LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-223485

The Complete Genome Sequence of the Marine, Chemolithoautotrophic, Ammonia-Oxidizing Bacterium Nitrosococcus oceani ATCC19707

Martin G. Klotz, Daniel J. Arp, Patrick S.G. Chain, Amal F. El-Sheikh, Loren J. Hauser, Norman G. Hommes, Frank W. Larimer, Stephanie A. Malfatti, Jeanette M. Norton, Amisha T. Poret-Peterson, Lisa M. Vergez, Bess B. Ward

August 7, 2006

Applied and Environmental Microbiology

Section: Evolutionary and Genomic Microbiology

Title: The Complete Genome Sequence of the Marine, Chemolithoautotrophic, Ammonia-Oxidizing Bacterium *Nitrosococcus oceani* ATCC19707

Running Title: *N. oceani* Genome Sequence

Authors: Martin G. Klotz[1]*, Daniel J. Arp[2], Patrick S.G. Chain[3,4], Amal F. El-Sheikh[1], Loren J. Hauser[5], Norman G. Hommes[2], Frank W. Larimer[5], Stephanie A. Malfatti[3,4], Jeanette M. Norton[6], Amisha T. Poret-Peterson[1], Lisa M. Vergez[3,4], and Bess B. Ward[7].

Addresses: University of Louisville, Louisville, KY 40292[1], Oregon State University, Corvallis, OR 97331[2], Lawrence Livermore National Laboratory, Livermore, CA 94550[3], Joint Genome Institute, Walnut Creek, CA 94598[4], Oak Ridge National Laboratory, Oak Ridge, TN 37831[5], Utah State University, Logan, UT 84322[6], and Princeton University, Princeton, NJ 08544 [7]

*Corresponding Author. Mailing address: Department of Biology, University of Louisville, 139 Life Science Building, Louisville, KY 40292. Phone: (502) 852-7779. Fax: (502) 852-0725. E-mail: aem@mgklotz.com

Keywords: nitrification, microbial genome, ammonia-oxidizing bacteria, *Nitrosococcus oceani*

ABSTRACT

The Gammaproteobacterium, *Nitrosococcus oceani* (ATCC 19707), is a Gram-negative obligate chemolithoautotroph capable of extracting energy and reducing power from the oxidation of ammonia to nitrite. Sequencing and annotation of the genome revealed a single circular chromosome (3,481,691 bp; 50.4% G+C) and a plasmid (40,420 bp) that contain 3052 and 41 candidate protein-encoding genes, respectively. The genes encoding proteins necessary for the function of known modes of lithotrophy and autotrophy were identified. In contrast to betaproteobacterial nitrifier genomes, the *N. oceani* genome contained two complete *rrn* operons. In contrast, only one copy of the genes needed to synthesize functional ammonia monooxygenase and hydroxylamine oxidoreductase, as well as the proteins that relay the extracted electrons to a terminal electron acceptor were identified. The *N. oceani* genome contained genes for 13 complete two-component systems. The genome also contained all the genes needed to reconstruct complete central pathways, the tricarboxylic acid cycle and the Embden-Meyerhof-Parnass and pentose phosphate pathways. The *N. oceani* genome contains the genes required to store and utilize energy from glycogen inclusion bodies and sucrose. Polyphosphate and pyrophosphate appear to be integrated in this bacterium's energy metabolism, stress tolerance and the ability to assimilate carbon via gluconeogenesis. One set of genes for type I RuBisCO was identified, while genes necessary for methanotrophy and for carboxysome formation were not identified. The *N. oceani* genome contains two copies each of the genes or operons necessary to assemble functional complexes I and IV as well as ATP synthase (one H^+-dependent F_0F_1-type, one Na^+-dependent V-type).

INTRODUCTION (→ Fig. 1. TEM of *N. oceani*)

The ammonia-oxidizing bacterium *Nitrosococcus oceani* ATTC 19707 (Fig. 1; Bacteria, Proteobacteria, Gammaproteobacteria, Chromatiales, *Chromatiaceae*, *Nitrosococcus*, *Nitrosococcus oceani*) was the first ammonia-oxidizing bacterium isolated by enrichment culture from seawater (*Nitrosocystis oceanus*, (53)) and it resembles the original type strain, *Nitrosococcus winogradskyi* 1892, which was lost. As a member of the order Chromatiales, the purple sulfur bacteria, *N. oceani* is a member of the evolutionarily oldest taxonomic group capable of lithotrophic ammonia catabolism. To date, *N. oceani* and *N. halophilus* are the only recognized species of gammaproteobacterial ammonia-oxidizing bacteria (AOB). All other cultivated aerobic AOB are Betaproteobacteria and their members have been detected in soils, freshwater and sediments as well as marine environments (40). In contrast, gammaproteobacterial AOB have only been found in marine or saline environments. *Nitrosococcus oceani* has been detected in many marine environments using immunofluorescence (72, 78), and more recently on the basis of cloned gene sequences from DNA extracted from natural seawater (61, 74). In addition to the truly marine environment, *N. oceani* was detected by immunofluorescence and fluorescent in situ hybridization (FISH) in the saline waters of Lake Bonney, a permanently ice-covered lake in Antarctica (70). *Nitrosococcus halophilus* has been isolated only from saline ponds (38) and has not been detected in other environments by using molecular probes.

The general role of nitrifying bacteria in marine systems is to link the oxidizing and reducing processes of the nitrogen cycle by converting ammonium to nitrate. This conversion is responsible for maintaining nitrate, the major component of the fixed nitrogen pool in the oceans,

which is present almost everywhere below a few hundred meters at concentrations approaching 40 μM. The deep nitrate reservoir of the oceans, believed to have come about by abiotic processes on the primordial Earth (48), is a huge pool of nitrogen whose availability to primary producers in the surface layer is still controlled largely by physical processes (27). The nitrification process produces oxidized forms of nitrogen that are lost via denitrification and anaerobic ammonia oxidation (anammox) (34, 65). By converting nitrogenous compounds released as waste products of metabolism into NO_x intermediates that can act both as oxidants and reductants for the fixed N removal processes (14, 73), nitrification closes the global nitrogen cycle. Nitrification occurs in both the water column and in sediments of marine environments. In sediments, nitrification is often tightly coupled with denitrification, and can account for a significant fraction of the total oxygen consumption in sediments (80). It has been discovered only recently, that aerobic ammonia-oxidation is also carried out by some Crenarchaeota ((58) and references therein). Könneke et al. (45) reported the isolation of a marine crenarchaeote, *Candidatus Nitrosopumilus maritima*, that was able to grow chemolithoautotrophically by aerobically consuming ammonia and producing nitrite. This physiological observation has been supported by the identification of DNA sequences similar to the genes encoding the three subunits of ammonia monooxygenase (AMO) from AOB (45). On the other hand, the recently completed genome sequence of the Crenarchaeote, *Cenarchaeum symbiosum*, a symbiont of marine sponges (28) predicted to be an Ammonia-Oxidizing Archaeon (AOA), lacked all open reading frames with sequence similarity to genes known to be essential to ammonia-oxidation in all AOB such as hydroxylamine oxidoreductase (HAO) and cytochromes c554 and cm552 (7). Furthermore, the *C. symbiosum* genome lacked the genes for other cytochromes proteins known to control nitrosating stress in AOB (31) except for one (*norQ*) of the four required genes for

functional NO-reductase (28). In contrast, the *C. symbiosum* genome contained numerous genes that encode putative copper blue proteins (54) including one (CENSYa_1582) with significant sequence similarity to the NcgA and NirK protein family, which is implicated in nitrite reduction by AOB (6). Since AMO activity is also copper-dependent, it appears that ammonia oxidation in Archaea is entirely based on the function of copper-containg proteins. Because copper is known to be redox-active only under oxic conditions and because the cytochrome-based core module of bacterial ammonia catabolism (encoded by the *hao* gene cluster) has evolved from bacterial inventory involved in (anaerobic) denitrification (7)), it is highly likely that ammonia-catabolism in the Archaea has evolved fairly late by incorporating an AMO-like function into an ammonia-independent metabolism. Recent phylogenetic analyses of known Amo subunit protein sequences ((28, 58) and M.G. Klotz, unpublished results) suggest, indeed, that the AOA likely obtained their *amo* genes via lateral transfer from AOB. It will thus be interesting to identify genetically and biochemically how the AOA resolved the tasks of aerobic ammonia oxidation and detoxification of the resulting NO_x compounds.

The susceptibility of ammonia oxidizers to inhibition by sunlight (due to the light sensitivity of ammonia monooxygenase; AMO) is probably responsible for the characteristic distribution of nitrification in the water column; maximal rates occur in surface waters near the bottom of the euphotic zone (74). Nitrification rates decrease with increasing depth as the rate of organic matter decomposition (and thus ammonium supply) decreases with depth. As a consequence of both nitrifier and denitrifier activities, the oceans emit large amounts of the greenhouse gas nitrous oxide (57). The total oceanic N_2O inventory is about 2/3 the size of the total atmospheric inventory and the oceanic N_2O flux to the atmosphere is estimated to be 4 Tg N/year (57). It has been shown that nitrous oxide and nitric oxide can be produced by aerobic

AOB either through the reduction of nitrite (NO_2^-; (17, 66)) or the oxidation of hydroxylamine (NH_2OH; (35)). Goreau et al. (25) reported on N_2O production by a marine *Nitrosomonas* isolate and a culture of *N. oceani* that had been isolated from the Western Atlantic by Stan Watson. N_2O production was found to be much higher under low than under atmospheric O_2 conditions, indicating that nitrifiers may produce significant amounts of N_2O in the interface between the oxic and anoxic zones. Recent geochemical evidence indicates that most of the N_2O in the ocean is derived from nitrification (56); therefore, the detailed understanding of the marine nitrification process is crucial for any global management strategy of greenhouse gas production. This significance of nitrification in the oceans and its indirect influence on the oceanic carbon budget led to the selection of *N. oceani* ATTC 19707 as a target for genome sequencing in the Genomes to Life microbial sequencing program of the US Department of Energy.

MATERIALS AND METHODS

Nitrosococcus oceani strain ATCC19707 was obtained from the American Type Culture Collection and maintained at a temperature of 30°C in the dark on marine medium as described previously (1, 39, 59). For the isolation of genomic DNA, cultures were grown in 0.6 and 1-L batches of medium in 2 and 4-L Erlenmeyer flasks, respectively, titrated to pH 8.0 daily with K_2CO_3.

Library construction, sequencing, and sequence assembly. Genomic DNA was isolated from late-exponential phase cultures of *N. oceani* strain ATCC19707 as described by McTavish *et al.* (49) modified following the recommendations by the Department of Energy's Joint Genome Institute (DOE-JGI, Walnut Creek, CA). The genome was sequenced using the

whole-genome shotgun method as previously described (16, 23). Briefly, random 3 and 8-kb DNA fragments were isolated and cloned into pUC18 and pMCL200 vectors, respectively, for amplification in *Escherichia coli*. A larger fosmid library was constructed containing approximately 40kb inserts of sheared genomic DNA cloned into the pCC1Fos cloning vector. Double-ended plasmid sequencing reactions were performed by the DOE JGI using ABI 3730xl DNA Analyzers and MegaBACE 4500 Genetic Analyzers as described on the JGI website http://www.jgi.doe.gov/.

After quality control of the 60,402 total initial reads of draft sequence, 51,334 were used for the final assembly, producing an average of 9.3 fold coverage across the genome. Processing of sequence traces, base calling and assessment of data quality were performed with PHRED and PHRAP, respectively. Assembled sequences were visualized with CONSED. The initial assembly consisted of 36 contigs. Gaps in the sequence were primarily closed by primer walking on gap-spanning library clones or with PCR products from genomic DNA. True physical gaps were closed by combinatorial (multiplex) PCR. Sequence finishing and polishing added 249 reads and assessment of final assembly quality was completed as described (16).

Sequence analysis and annotation. Automated gene modeling was completed by combining results from Critica, Generation, and Glimmer modeling packages, and comparing the translations to GenBank's nonredundant (NR) database using basic local alignment search tool for proteins (BLASTP). The protein set was also searched against KEGG Genes, InterPro, TIGRFams, PROSITE, and Clusters of Orthologous Groups of proteins (COGs) databases to further assess function. Manual functional assignments were assessed on individual gene-by-gene basis as needed.

Nucleotide sequence accession number. The sequence and annotation of the complete *N. oceani* strain ATCC19707 genome is available at GenBank/EMBL/DDBJ accession numbers CP000127 (chromosome) and NC_007483 (plasmid).

RESULTS AND DISCUSSION

General features **(→ Table 1. Genome features; Figure 2. Circular representation of the genome (chromosome and plasmid maps).**

Genome properties: The *N. oceani* ATCC19707 genome is comprised of a single circular chromosome (3,481,691 bp; 50.4% G+C) and a plasmid (40,420 bp) that contain 3052 and 41 candidate protein-encoding genes, respectively (Fig. 2). Most (91%) of candidate genes were in orthologous clusters ORFs of published genomes and a total of 76.5%, 68.7% and 64.9% had hits with ORFs in the COG, Pfam and InterPro data bases, respectively (Tab. 1). The majority of genes (66.9%) could be assigned a function; however, only 11.4% of these genes were assigned to enzymes and only 8.6% were connected to the KEGG pathways. In contrast, 23.8% of the genes in the *N. europaea* genome were assigned to enzymes and 18.6% were connected to the KEGG pathways (http://img.jgi.doe.gov, (16)). The taxonomic breakdown of best BLASTP hits against the KEGG completed genomes database, is as follows: Gammaproteobacteria (1424 genes) followed by Betaproteobacteria (521), Cyanobacteria (183), Alphaproteobacteria (112) and Delta-/Epsilonproteobacteria (121). Individual top hits were with ORFs from *Methylococcus capsulatus* (545) followed by *Nitrosomonas europaea* (224), *Pseudomonas aeruginosa* PAO1 (189), Azoarcus sp. EbN1 (161), *Pseudomonas putida* (75), *Pseudomonas syringae* (74), and *Geobacter sulfurreducens* (69).

183 In contrast to betaproteobacterial nitrifier genomes, the *N. oceani* genome contained two
184 complete *rrn* operons that belong to different classes. These operons are located on different
185 replicores, neither of which was near the origin of replication. The *rrn* operon on the plus strand
186 belongs to the class that contains Ala-tRNA and Ile-tRNA genes between the 16S and 23S genes,
187 whereas the *rrn* operon on the minus strand has no inserted genes. While the 16S-23S intergenic
188 region in rrn1 contains 714 bp, including the two tRNAs, the 16S-23S intergenic region in rrn2 is
189 only 224 bp – of which 45 bp downstream of the 16S and 107 bases upstream are identical to
190 those positions in rrn1. However, the16S, 23S and 5S rRNA genes themselves are 100%
191 identical.

192 The plasmid comprises mostly hypothetical and conserved hypothetical proteins. A
193 transposase (Noc_A0021), phage integrase (Noc_A0015) and a small number of other phage
194 related genes along with restriction modification systems are also found on this replicon. A
195 possible replication protein (Noc_A0039), whose only putative homologue (56% similar over
196 80% of its length) is the RepA protein encoded in plasmid pRA2 from *Pseudomonas alcaligenes*
197 (43), together with a possible partitioning system (Noc_A0013-A0014) may help this plasmid be
198 maintained within the *N. oceani* population.

199 **Families and clusters of foreign and repeat sequences.** Surprisingly, we identified
200 several large blocks of genes that were identified as putatively phage-related, indicating that *N.*
201 *oceani* has been a frequent target for bacteriophages in the open ocean. Ten regions, ranging in
202 size from 6.4 to 44.3-kb (totaling >175 kb), were found to harbor bacteriophage remains, such as
203 phage integrase, terminase, primase, and tail genes. These regions were found to be associated
204 with tRNAs (flanking), restriction modification systems, transposases, virulence-associated and

205 many hypothetical genes. Other phage hits are not associated with any cluster of genes, but are
206 distributed throughout the genome.

207 There are five families of identical or nearly identical (>95%) copies of insertion
208 sequence (IS)-like elements that encode putative transposases, the largest family of which
209 comprises 15 copies of an IS that encodes two transposases. Two of the families, with three
210 members each, share ~83% identity and represent an IS superfamily. Interestingly, the genome
211 encodes many predicted transposases (127), some of which cluster in large groups, the largest of
212 which comprises 15 consecutive ORFs (Noc_0003-0017) encompassing >7 kb. Also, of the 14
213 predicted phage integrase genes or pseudogenes, two were nearly identical frameshifted integrase
214 genes (Noc_0080, Noc_1095).

215 Two identical copies of the Tu translation elongation factor (Noc_2326, Noc_2338) and
216 two gene copies encoding a fatty acid desaturase, a metalloprotein were identified. Additionally,
217 there are two copies of a gene annotated as ammonia permease (Noc2700 and Noc2701);
218 however, both were much shorter than AmtB proteins from other organisms and aligned only
219 with their C-termini. Furthermore, the region containing the membrane-spanning domains lacked
220 amino acid residues that are conserved in other AmtB proteins. Hence, operation of AmtB-
221 facilitated ammonia uptake by one or both products of the Noc_2700 and Noc_2701 genes needs
222 to be experimentally verified. In addition to these examples, we have found several lines of
223 evidence suggesting that genes or gene families are undergoing duplication and diversification.
224 Examples include Noc_1310 which appears to be a truncated C-terminal version of Noc_1552, a
225 full-length phosphoenolpyruvate-protein phosphotransferase, and the gene Noc_0725, which is
226 only ~70% identical to one of three nearly identical copies (Noc_0343, Noc_0724, Noc_0973) of
227 a predicted hypothetical protein of 492 amino acids.

Information processing and modification systems: The genome of *N. oceani* contains the complete sets of genes necessary to encode DNA-directed DNA polymerases I (Noc_0554) and III (Noc_0002, Noc_0288, Noc_0846, Noc_1659, Noc_2593, Noc_2663, Noc_2814) as well as multiple copies of the gene encoding the epsilon subunit of Pol III. Genes encoding Pol II were not identified. *Nitrosococcus oceani* is equipped with a full complement of genes to carry out repair of DNA lesions (*uvrABCD*, *recFRO*, *mutSHLT*, *radC*, *recN*) and recombination (*recA*, *recD* and *recG*; *ruvABC*), which may be necessary as a consequence of exposure to mutagens and uptake of foreign DNA. The genome of *N. oceani* contains a high number of open reading frames (a total of 24) that code for type I and type III site-specific restriction endonucleases (REs) as concluded from a comparison with available complete genome sequences of ammonia-oxidizing Betaproteobacteria and closely related Gammaproteobacteria whereof the next highest number of ORFs was found in the genome of *Xylella fastidosa* Ann-1 (a total of 13). The presumed function of these restriction modification systems is protection against phage infection and foreign DNA by recognizing specific methylation patterns and distinguishing between host and foreign DNA (11, 15). The reason for this high number of REs in the genome of *N. oceani* is presently unclear; however, a similarly high number of ORFs encoding site-specific endonucleases have been identified in the yet unfinished genome of *Nitrosococcus halophilus* (A.F. El-Sheikh and M.G. Klotz, unpublished results).

The genes encoding the subunits of DNA-directed core RNA polymerase (EC:2.7.7.6) did not reside in a single gene cluster in the *N. oceani* genome. Whereas the genes encoding the beta and beta' subunits were arranged in tandem (Noc_2331 and Noc_2330, respectively), the gene encoding the alpha subunit resided upstream of this tandem as gene Noc_2300. A gene encoding an omega subunit (COG1758) was also found (Noc_1213). The genome also contained

a variety of genes that encode alternative sigma factors, some of which were found in multiple copies (see below).

Metabolism and Transport: Genomic Basis of Ammonia-Lithotrophy

Energy metabolism - ***Acquisition of reductant from the environment****:* A cluster of three contiguous genes encoding the subunits of ammonia monooxygenase (Noc_2503-2501) were found to be organized in overlapping operons as described previously (1, 59). No additional functional *amo* genes or *amo* pseudogenes were found in the genome. As reported previously, the terminator of the *amo* operon was succeeded by a transcriptional unit containing the *orf5* gene (Noc_2500) (59). An additional *orf5*-like gene was found as an orphan in the genome (Noc_3006). The deduced Orf5 protein sequences revealed a signal peptide and an additional membrane-spanning domain at its C-terminus. Such *orf5* genes with a high degree in sequence identity and in conservation of synteny have also been identified in betaproteobacterial nitrifier genomes downstream of the *amo* operon (U92432, AF016003); however, as one of two sequence-related genes in an *orf45* transcriptional unit (16). Interestingly, an *orf5* homologue was also found in the whole genome sequence of the gammaproteobacterial methanotroph, *Methylococcus capsulatus* (Bath) (MCA2130), where it was not in proximity of either of the two gene clusters that encode particulate methane monooxygenase (pMMO), a homologue of AMO (29, 59). Instead, the *orf5* gene was resident in a transcriptional unit together with a gene that encodes a *pan1*-type multi-copper oxidase (MCO; MCA2129), which is likely a homologue to the pan1-type MCO-encoding gene found upstream of the *amo* operon in the *N. oceani* genome (Noc_2506). Because both AMO and pMMO can oxidize ammonia to hydroxylamine and because of the sequence and genome organizational similarities between the associated *mco* and

274 *orf5* genes in the *N. oceani* and *M. capsulatus* genomes, a functional role of their membrane-
275 associated expression products in transfer of electrons or intermediates related to ammonia
276 oxidation is proposed.

277 The oxidation of hydroxylamine is the core of ammonia catabolism as it provides
278 electrons for redox-dependent proton pumping (cytochrome bc_1 complex, terminal cytochrome *c*
279 oxidase). Cytochrome P460 (Noc_0890) is a likely ancient hydroxylamine (and nitric oxide)
280 dehydrogenase suited to detoxify the mutagenic hydroxylamine at low turnover rates (9, 10).
281 However, the electrons gained during the oxidation process cannot be transferred directly to the
282 cytochrome *bc1* complex in the membrane and are likely relayed via the periplasmic soluble
283 cytochrome *c*552 (Noc_0751) to the terminal oxidase. While this pathway contributes to the
284 proton motive force, it is an inefficient use of reductant. A more efficient pathway occurs when
285 hydroxylamine is catalytically oxidized by hydroxylamine oxidoreductase (HAO, Noc_0892)
286 and electrons are funneled via cytochromes *c*554 (Noc_0894) and *c*m552 (Noc_0895) to the
287 ubiquinone pool (Q/QH_2, Noc_1248-1252). The structure and sequence conservation of the
288 HAO-c554-cm552 pathway in *N. oceani* was recently evaluated (7), based in part on the whole
289 genome sequence reported in this paper. The reduced quinone pool provides reductant to the
290 AMO complex, to the cytochrome bc_1 proton-pumping complex (Noc_0297–0299) in the
291 electron transport chain, and to the NUO complex I responsible for reverse electron transport
292 needed to generate NADH (see below). In addition, numerous uncharacterized MCO-encoding
293 genes were found in the genome (Noc_0889 and Noc_2605, Type-1 MCOs; Noc_1542, exported
294 MCO; Noc_1741, copper resistance protein), which need further characterization to assess their
295 involvement in catalytic and electron transfer processes of catabolic pathways in *N. oceani*. In
296 addition, the finding of a gene encoding the red copper protein, Nitrosocyanin (Noc_1090), in

297 the *N. oceani* genome suggests that this protein is unique to and important for ammonia
298 catabolism.

299 The hydrolysis of urea to ammonia and carbon dioxide can be carried out in *N. oceani* by
300 the ATP-independent hetero-multimeric nickel enzyme urea-amidohydrolase (urease) (39).
301 Ureolysis could be beneficial to the cell because it produces *N. oceani*'s sole sources for energy
302 and reductant (ammonia) and carbon (CO_2). Given the low concentration of dissolved urea in the
303 oceans, the role of urease in this organism is unclear.

304 Some betaproteobacterial nitrifiers can utilize H_2 (12). However, the genome of *N.*
305 *oceani* did not contain any genes that encode subunits of a hydrogenase, which is in agreement
306 with the fact that *Nitrosococcus* cannot grow on H_2 as the sole source of energy and reductant.

307 **Energy metabolism -** ***Electron flow, generation of universal reductant, generation of***
308 ***the proton gradient and ATP production*****.** The gene profile for *N. oceani* reveals complete sets
309 of genes for electron transfer from NADH to O_2 via NADH quinone oxidoreductase (Complex
310 I), cytochrome bc_1 complex (Complex III) and a Cu-aa_3 type cytochrome *c* oxidase (Complex
311 IV). Coupled with genes for a complete tricarboxylic acid (TCA) cycle and glycolytic pathway,
312 it seems that *N. oceani* has the potential to gain energy through the oxidation of organic
313 compounds. A complete pentose phosphate pathway provides an alternative mechanism for
314 oxidizing sugars and generating NAD(P)H. It appears that the obstacle to an organotrophic
315 mode of catabolism is *N. oceani*'s inability to import suitable organic substrates. When using
316 ammonia as the energy source (lithotrophy), there is a need to generate NAD(P)H through
317 reverse electron flow and a complete TCA cycle is not needed. The complexes involved in
318 electron flow, generation of reductant and the proton gradient, and production of ATP are
319 described.

Two complete yet different sets of genes encoding Complex I (NDH-1/NADH Quinone oxidoreductase) are present in the genome of *N. oceani*. Genes Noc_1115-1127, which encode one copy of Complex I, are most similar to Complex I genes found in other Gammaproteobacteria. This operon contains only 13 genes; the c and d subunits are fused into a single gene. The second set of genes encoding a Complex I (Noc_2552-2565) includes genes with top blast hits to *N. europaea* (6 genes), to other Betaproteobacteria (3 genes), and to Gammaproteobacteria (5 genes). The role of these distinct complexes in *N. oceani* is unknown. It may be that the complex with the strongest similarity to that found in *N. europaea* plays a role in reverse electron flow when ammonia is the sole electron donor, whereas the gammaproteobacterial complex may be important in forward electron flow associated with NADH oxidation. The *N. oceani* genome also contains the genes needed to encode a Na^+-transporting NADH:ubiquinone oxidoreductase (Noc_0970, Noc_1170-1174).

The presence of candidate genes for a Na^+-dependent complex I (Noc_0970, Noc_1170-1174), a Na^+-dependent V-type ATPase (Noc_2081-2089), and several Na^+-/H^+ antiporters (Noc_0159, 0521, 1282, 2134, 2952) raises the possibility of a sodium circuit in addition to the proton circuit in *N. oceani*, which is likely an adaptation to *N. oceani's* high salt environment. Under chemolithotrophic conditions, reverse operation of the sodium-dependent Complex I in the plasma membrane could generate additional NADH; however, the sodium-dependent ATPase would be needed to remove excess sodium from the cytoplasm. Under mixotrophic conditions, whether lithoheterotroph or organoautotroph, a sodium gradient might help provide at least some of the organic needs of the cell through import (4). Our finding of several sodium-dependent transporters in the genome (Noc_0779, 1365, 1575, 1600, 2446, 2711) support this possibility; however, operation of the sodium-dependent ATPase would, again, be needed to

remove excess sodium from the cytoplasm. If import of certain organics enable *N. oceani* of organotrophy, additional NADH could be generated. The additional NADH produced by the sodium circuit (Na^+-dependent complex I and transporters) on the expense of ATP generated by the proton circuit, could serve the "forward" complex I as an extension of the bacterium's electron transport chain thereby effectively converting a sodium-motive force into a proton-motive force. This could be of utility for chemotaxis and proton-dependent transport.

The genes encoding a ubiquinol-cytochrome *c* reductase (the cytochrome bc_1 complex) are located in an operon (Noc_0297-0299). *Nitrosococcus oceani* has genes encoding two complete terminal cytochrome c oxidases (COX), both of the Cu-aa_3 type. One complex (Noc_3044-3047) appears to be of gamma- and betaproteobacterial descent (*Pseudomonas/ Nitrosomonas*) and the other (Noc_1244--1247) of Bacteroidetes/Chlorobi decent. There are two additional copies of the genes for subunit 1 and subunit 2, but not for subunit 3. In addition, there are eight copies of genes encoding Class I *c*-type cytochromes.

Additional electron sinks are also encoded in the genome, i.e., there are three di-heme cytochrome *c* peroxidases (Noc_0488, Noc_1263, Noc2697). A partial denitrification pathway is present, including a Cu-type nitrite reductase (*nirK*; Noc_0089) and nitric oxide reductase (*nor*; Noc_1847-1851). However, genes for nitrate reductase (Nar) and nitrous oxide reductase (Nos) were not identified.

Under lithotrophic conditions, the proton gradient is presumably generated by the action of the quinol-cytochrome *c* oxidoreductase and cytochrome *c* oxidase. Under organotrophic conditions, we assume that one or both of the NADH-ubiquinone oxidoreductases could also contribute to the generation of the proton gradient. In addition, gene Noc_1901 encodes a proton-translocating inorganic pyrophosphatase, which, in the presence of a flux of pyrophosphate,

366 could contribute to the generation of a proton gradient (Fig. 3). Alternatively, the enzyme could
367 use the proton gradient to generate pyrophosphate necessary for glycolysis and other processes.
368 A potential source of pyrophosphate could be polyphosphate since a polyphosphate kinase is
369 present (Noc_2388). ATP formed from degradation of polyphosphate could release
370 pyrophosphate through the action of one of the many nucleoside diphosphate hydrolases
371 (NUDIX hydrolases) encoded in the genome (Noc_0193, Noc_0306, Noc_2018, Noc_2420,
372 Noc_2512, Noc_2643, Noc_2749).

373 A typical proteobacterial H^+-translocating FoF1-type ATP synthase is encoded in an
374 operon (Noc_3073-3080). Additionally, the genome also encodes a bacterial V-type ATP
375 synthase (Noc_2081-2089). Subunits A through I, with the exception of G, are encoded in this
376 operon. A gene encoding subunit G does not appear to be present in the *N. oceani* genome;
377 however, subunit G is not present in all bacterial V-ATPases (46). It appears to function with
378 subunit D in formation of a peripheral stalk. Genes for subunits A and B, which form the active
379 site of the enzyme, are highly conserved while the remaining genes in the cluster are less well
380 conserved.

381 **Energy storage strategies - *Sucrose Synthase, Sucrose-Phosphate Synthase***:
382 *Nitrosococcus oceani* has genes encoding a sucrose synthase (SuSy; Noc_3068) and sucrose
383 phosphate synthase (SPS; Noc_3069). Sucrose synthase in plants functions predominantly in the
384 degradation of sucrose (47). Sucrose phosphate synthase functions in plants to synthesize
385 sucrose phosphate and sucrose phosphate phosphatase completes the synthesis of sucrose in
386 plants. A separate gene for sucrose phosphate phosphatase is not present in the genome of *N.*
387 *oceani*; however, gene Noc_3069 is a fusion of the genes for sucrose phosphate synthase and
388 sucrose phosphate phosphatase. The HAD triad conserved in this superfamily of

389 glycohydrolases is encoded in gene Noc_3069. UDP-glucose is the glucosyl donor for both
390 enzymes, and gene Noc_2280 encodes a UTP glucose-1-phosphate uridylyltransferase. Gene
391 Noc_3067 encodes a fructokinase that could make the fructose-6-phosphate required by sucrose
392 phosphate synthase.

393 Sucrose synthesis in bacteria is not well documented or understood (47). Sucrose and
394 sucrose synthesizing activities were detected in two species of halotolerant methanotrophs,
395 *Methylobacter alcaliphilus* 20Z and *Methylobacter modestohalophilus* 10S (37). Genes for
396 sucrose synthesis have been identified, so far, in just a few Proteobacteria, including
397 *Acidithiobacillus ferrooxidans* and *N. europaea*. The genes are also present in cyanobacteria
398 where their function in sucrose synthesis has been demonstrated. Their role in *N. oceani* is
399 unknown. Perhaps sucrose can serve as an osmoprotectant, as has been suggested for other
400 prokaryotes.

401 **Energy storage strategies - *Carbon Storage Products*:** The genome of *N. oceani* also
402 contains genes encoding the five enzymes necessary for the synthesis of glycogen from fructose-
403 6-phosphate. The genes encoding glucose-1-phosphate adenyl transferase (Noc_0905) and 1,4-
404 alpha glucan branching enzyme (Noc_0904) are contiguous. The gene for phosphoglucomutase
405 was identified in *N. oceani* (Noc_1719) but genes necessary for the synthesis of poly beta-
406 hydroxybutyrate do not appear to be present in the genome of *N. oceani*.

407 **Central carbon metabolism:** The gene profiles are consistent with complete pathways
408 for glycolysis (from phosphorylated sugars or glucose) and gluconeogenesis. However, the
409 mechanism for the interconversion of fructose-6-phosphate and fructose-1,6-bisphosphate is not
410 clear. Gene Noc_0021 encodes fructose-1,6-bisphosphatase. A candidate for an ATP-dependent
411 phosphofructokinase is not present, but gene Noc_2846 shows some similarity to genes encoding

412 diphosphate-fructose-6-phosphate 1-phosphotransferase (EC:2.7.1.90, pyrophosphate-dependent
413 phosphofructokinase). Both a membrane-bound, proton-translocating pyrophosphatase
414 (Noc_1901) and soluble pyrophosphatase (Noc_1134) are encoded, providing a mechanism to
415 hydrolyze pyrophosphate generated during gluconeogenesis. The proton-translocating enzyme
416 could also use energy from the proton gradient to synthesize pyrophosphate necessary for
417 glycolysis.

418 The genes encoding a complete pentose phosphate pathway are also present. We cannot
419 determine if glucose-6-phosphate dehydrogenase would couple to $NADP^+$ (as is most often the
420 case) or NAD^+ (which occurs less frequently). A transhydrogenase, encoded by genes Noc_0261
421 and Noc_0262, would allow transfer of reductant from NADH to $NADP^+$ and from NADPH to
422 NAD^+.

423 All the genes necessary for a complete TCA cycle are present. Of particular interest was
424 the presence of genes encoding alpha-ketoglutarate dehydrogenase because this enzyme activity
425 was missing in several obligate chemolithotrophs, including *N. europaea* (30). The absence of
426 alpha-ketoglutarate dehydrogenase activity would create a break in the TCA cycle that would
427 preclude organotrophy; however, if present, the role of alpha-ketoglutarate dehydrogenase in
428 these exclusively or predominantly lithotrophic organisms is not known. As was the case for *N.*
429 *europaea* (16), the genes encoding the three subunits (E1-3) of alpha-ketoglutarate
430 dehydrogenase were identified in the *N. oceani* genome and they were contiguous (Noc_0111-
431 0113). Because pyruvate dehydrogenase catalyzes a reaction mechanistically similar to that of
432 alpha-ketoglutarate dehydrogenase and has a similar subunit structure, one gene (*lpd*) encodes
433 the E3 subunit for both enzymes in many organisms. However, in *N. oceani*, pyruvate

dehydrogenase is encoded by separate genes, Noc_1254-1256, which includes the gene for subunit E3.

Amino acid and nucleotide metabolism: The *N. oceani* genome contains genes encoding the biosynthesis of the twenty amino acids required for the synthesis of proteins. Unless otherwise indicated, all amino acids mentioned were the L-form. The genes identified indicate most pathways are similar to previously identified synthesis pathways (5, 26, 68, 69, 77) and the genes encoding the enzymes have high similarity values with those found in other Proteobacteria, most of them with best matches to *Methylococcus capsulatus* or within the pseudomonads. While most biosynthetic pathway elements have been identified, missing enzymatic steps are typically involved with dual function enzymes that may be difficult to identify based on sequence alone (for example Noc_0176 encodes 3-phosphoshikimate 1-carboxyvinyltransferase with dual functions (EC:1.3.1.1 and EC:2.5.1.19)). In contrast, few amino acid degradative enzymes (catabolic enzymes) or specific transport genes were identified. Scavenging from leucine, valine and isoleucine may be possible. Several transaminases were identified but most could not be assigned to specific amino acids. Special attention was paid to amino acid biosynthetic functions found in the last common ancestor as indicated by their universal distribution in the three domains of life (44).

Several large amino acid biosynthetic operons were identified including 1) a mixed function supraoperon similar to that found in *Pseudomonas* encoding aromatic amino acid biosynthesis (76) and the *his* operon (22). The aromatic supraoperon is found encoded by Noc_0172-0177. The dual-function chorismate mutase/prephenate dehydratase P-protein is encoded by gene Noc_0174. The histidine biosynthetic genes are not in a single cluster and are located in the genome as genes Noc_2778-2779 (*hisDG*) and Noc_3051-3057 (*hisCBHAFIE* and

457 *hitA*). Split organizations of the *his* operon are relatively common and were also found in *N.*
458 *europaea* (16, 22). The genes *hisI* and *hisE* overlap by 8 bp but do not appear to be fused. The
459 gene encoding histidinol phosphatase (EC:3.1.3.15, Noc_0374) was found outside of the operon.
460 In the genome of *N. oceani* 20 aminoacyl-tRNA synthetases (AARS) were identified,
461 including two forms of LysRS (Class I and II) and, two distinct GlxRS (EC:6.1.1.17), but
462 missing AsnRS (EC:6.1.1.22) and GlnRS (EC:6.1.1.18). The most common organism for the top
463 match for the AARS was *Methylococcus capsulatus*, as is the case for the overall taxonomic
464 distribution of top matches. The Class II aminoacyl-tRNA synthetases for Phe and Gly have two
465 non-identical subunits and genes for both the alpha and beta subunits have been identified
466 adjacent to each other in the genome. While no AARS was identified as the specific GlnRS-type
467 (EC:6.1.1.18), it is likely that this function is mediated by the product of genes Noc_0264 or
468 Noc_2250, the non-discriminatory type GlxRS. Genes encoding a possible glutamyl-tRNA-Gln
469 amidotransferase (*gatCAB*) were found in the genome although *gatB* (Noc_2014) is separated
470 from *gatCA* (Noc_2635-2636). So an indirect route for synthesis of glutaminyl-tRNA is
471 probable. The sequence data alone are insufficient for designating either of glx-RS genes
472 specifically a discriminatory role. The presence of two genes does not appear to be a recent
473 duplication event as the peptides are more similar to their homologues in *Methylococcus* than to
474 each other (only 44% identical to each other over 315 aligned residues versus 55 and 67%
475 identical to the *Methylococcus* peptides).
476 No candidate for AsnRS was identified in the genome. The AspRS encoded by gene
477 Noc_0302 does contain the GAD domain typically found in AspRS involved in the indirect
478 transamination route to Asn—tRNA-^{ASN} synthesis. The *gatCAB* encoded transamination function
479 has a likely role in this pathway. Interestingly, the genome contains several distinct asparagine

480 synthetase genes of the glutamine dependent *asnB* type (Noc_0777, Noc_1965, Noc_1975,
481 Noc_2478). It is somewhat surprising but not without precedent (51) that multiple *asnB* genes
482 are present in an organism without an identified *asnRS*.

483 Although typically individual organisms contain only one class of LysRS, genes
484 encoding both classes of LysRS have been identified in the genome of *N. oceani*. This case of
485 LysRS existing in both Class I and Class II forms in the genome of *N. oceani* is the only known
486 bacterial exception to the AARS 'class rule' (60). The only known examples of the presence of
487 both classes of LysRS in a single genome are in the Archaea, *Methanosarcinia barkeri* and *M.*
488 *acetivorans.* In *N. oceani*, the class I LysRS is encoded by gene Noc_1618 with the best match to
489 the *M. barkeri* LysRS. Both selective retention and horizontal gene transfer have played roles in
490 the distribution of Class I LysRS in bacteria (2). The gene Noc_2625 encoding 356 amino acids
491 represents the C-terminal region of Class II LysRS with an intact core domain (best match is to
492 *Pseudoalteromonas haloplanktis* TAC125 (CAI85572)). Several other Gammproteobacteria have
493 a similar protein as described in COG 2269. In *N. oceani* this combination may be an example of
494 gene displacement of the LysRS Class II by a functional LysRS Class I gene (33). Alternatively,
495 in *M. barkeri* both forms are involved in the incorporation of the rare amino acid, pyrrolysine
496 into the enzyme monomethylamine methyltransferase. This is the first report of the presence of
497 Class I and Class II LysRS genes together in a bacterial genome (62).

498 Genes for the synthesis of all five purine and pyrimidine nucleotides are present. In
499 contrast, genes for degradation are very limited. There appears to be no capacity to catabolize
500 nucleotides with the exception of uridine, which can only be converted to pseudouridine.

501 **Transport:** Approximately 263 ORFs in *N. oceani*, about 9% of the total, are involved
502 with transport. Included are *P-P*-bond-hydrolysis-driven transporters, electrochemical-potential-

503 driven transporters, and channels/porins from a large number of protein families (Table 3).
504 Predominant among these are 23 ATP Binding Cassette (ABC) type transporters (85 genes) for a
505 variety of organic and inorganic substrates, the Resistance-Nodulation-Cell Division (RND)
506 family (26 genes) of H^+-antiport driven efflux transporters and a large number of genes (22)
507 involved with iron transport.

508 The *N. oceani* annotation lists at least 22 genes involved with iron transport. Iron
509 transport is particularly important due to the number of hemes integral to hydroxylamine
510 oxidoreductase and other cytochromes in the energy generating NH_3 oxidation pathway. Eleven
511 TonB-dependent iron siderophore receptors were identified (Noc_0321, Noc_0323, Noc_0326,
512 Noc_0541, Noc_0859, Noc_1269, Noc_1430, Noc_1489, Noc_1820, Noc_1925, Noc_2872)
513 including two TonB-dependent receptors for ferrienterochelin (Noc_0859, Noc_1269) and two
514 for ferrichrome (Noc_0321, Noc_0323). Unlike *N. europaea,* where many TonB-dependent
515 receptors are adjacent to genes encoding FecIR two-component sensor/regulatory proteins
516 involved with iron uptake, this was not the case for the TonB-dependent receptors in *N. oceani*.
517 No homologues of FecI or FecR were identified in the genome. However, two genes encoding
518 the ferric uptake regulator (Fur) were present (Noc_1194, Noc_2424) which may serve to
519 regulate the iron siderophore receptor expression. *Nitrosococcus oceani* has genes for the
520 synthesis of the hydroxamate-type siderophore aerobactin (Noc_1811-1814) and an aerobactin
521 receptor (Noc_1820). Additional iron transporters in *N. oceani* included an ABC-type
522 Fe^{3+}/cobalamine siderophore transport system (Noc_0838-0840) and a high affinity Fe^{2+}/Pb^{2+}
523 transporter (Noc_0164). While Fe^{2+} may not be abundant in marine environments, it may be
524 formed in the periplasmic space by a multicopper oxidase. The components of the
525 TonB/ExbB/ExbD type membrane energy-transducing complex were present in multiple copies

526 (TonB: Noc_0369, Noc_0610; ExbB: Noc_0142, Noc_0607-0608, Noc_2673; ExbD: Noc_0143,
527 Noc_0609, Noc_2674). No receptors for Fe^{+3}–citrate or Fe^{+3}-coprogen were found. In contrast
528 to *N. oceani*, the betaproteobacterial AOB *N. europaea* has over 100 genes involved in iron
529 transport including genes for receptors for Fe^{+3} coprogen but does not synthesize any
530 siderophores (16).

531 In addition to iron transport, a number of uptake systems (67 genes) for other inorganic
532 ions were identified. Inorganic N may be imported either as nitrite via a formate-nitrite (FNT
533 family) transporter (Noc_0109) or via a putative ammonia/ammonium permease (Amt)
534 (Noc_2700-2701). Sulfate could be imported via a sulfate permease (SulP family) (Noc_1626)
535 which may function as a SO_4^{2-}:H^+ symport or a SO_4^{2-}:HCO_3-antiport, or by a sodium:sulfate
536 symporter (Noc_1175). Phosphate transporters included two ABC transporters (Noc_2396-2399,
537 Noc_0581-0584) and two phosphate-selective porins (Noc_2417-2418) (see below). A number
538 of metal ion uptake and efflux systems were identified including an ABC transporter system
539 dedicated to Mn^{2+}/Zn^{2+} transport (Noc_2421-2423), a CorABC-type Mg^{2+}/Co^{2+} ion channel of
540 the MIT family (Noc_0240, Noc_1416, Noc_2263), three divalent cation transporters for
541 Mg^{2+}/Co^{2+}/Ni^{2+} of the MgtE family of magnesium transporters (Noc_1840, Noc_1785,
542 Noc_2801), and three divalent heavy-metal cation transporters (Noc_0092, Noc_04234,
543 Noc_1342). Efflux systems include two Small Multidrug Resistance (SMR) family transporters
544 of cations and cationic drugs (Noc_0601, Noc_2235), CopCD copper export proteins
545 (Noc_1741-1742), and four CDF-type cation efflux proteins (Co^{2+}/Zn^{2+}/Cd^{2+}) (Noc_0595,
546 Noc_1534, Noc_1782, Noc_2871). Other ion transporters include a chloride channel protein
547 EriC (Noc_0358), and an MFS-type cyanate transporter (Noc_1456). Two P-type cation
548 transporting ATPases were found (Noc_1406, N0c_2130).

549 Likely of importance to *N. oceani*'s salty habitat in the oceans are a number of
550 monovalent cation transporters. Several Na^+/H^+ antiporter systems were found including a
551 NhaD-type (Noc_1492), a NhaC-type (Noc_2134), and two NhaP-type (Noc_0159, Noc_0521).
552 These play important roles in maintaining intracellular pH and conferring salt-tolerance. Five
553 genes of the Small Conductance Mechano-sensitive Ion Channel (MscS) family (Noc_0108,
554 Noc_0602, Noc_1030, Noc_1853, Noc_1914) were found which play a role in fast
555 osmoregulatory responses. Other transport systems identified included the genes encoding both
556 the NAD-binding component (Noc_3020, Noc_3064) and the membrane component (Noc_0960,
557 Noc_2194, Noc_2952) of the KefB-type K^+ transport systems, a Trk-type K^+ uptake system
558 (Noc_0242, Noc_1639-1640), and two DASS family divalent anion/Na^+ symporters (Noc_1175,
559 Noc_2446).

560 The genome of *N. oceani* revealed only a few transporters for importing organic
561 compounds. ABC-type transporters may import dipeptides/oligopeptides (Noc_1344, Noc_1768,
562 Noc_2770-2773), proline/glycine betaine (Noc_0539-0540), and unspecified sugars (Noc_0279-
563 0282). Other transporters included an amino acid transporter (Noc_3063), a lactate permease
564 (Noc_1578), and urea transporter (Noc_2884). A number of potential transport systems for
565 dicarboxylates were found including a TRAP-type C4-dicarboxylate permease (Noc_0598,
566 Noc_0709-0710), a Tellurite-resistance/Dicarboxylate Transporter (TDT) (Noc_0077,
567 Noc_0542) which may be involved in transporting dicarboxylic acid intermediates, and a
568 DAACS family di- tri-carboxylate/amino acid:cation symporter (Noc_1175, Noc_2446).

569 Export systems for organic and toxic compounds included ABC transporters for organic
570 solvents (Noc_1746, Noc_2096, Noc_2782-2785), multidrug resistance (Noc_1779-1781,
571 Noc_2141-2142, Noc_1835-1836, Noc_2644-2646, Noc_2832-2833), heme export (Noc_0946-

572 0948), dipeptides (Noc_0933-0936, Noc_2184-2185), polar amino acids (Noc_1558-1559),
573 MFS-type arabinose efflux transporters (Noc_1547, Noc_1759, Noc_2803, Noc_3022), colicin
574 export (Noc_0144-0145, Noc_2616), polysaccharide/polyol phosphate (Noc_1227-1228,
575 Noc_2181-2182), lipoprotein export (Noc_1272-1273, Noc_2214-2216, Noc_2669-2670), as
576 well as a transporter involved in lysophospholipase L1 biosynthesis (Noc_0427-0428). Unlike
577 the uptake ABC transporters, these ABC transporters generally lacked any associated
578 periplasmic component.

579 *Nitrosococcus oceani* possesses several protein export and secretion systems including a
580 preprotein translocase (Noc_2305), genes for exporting folded redox proteins via the sec-
581 independent protein secretion system (TatABC, Noc_3058-3060), genes encoding Type II
582 general secretion/pilus synthesis pathway, and genes encoding the Type IV conjugal
583 DNA/protein transfer system.

584 *Nitrosococcus oceani* has a cluster of genes encoding parts of a PTS-type sugar transport
585 system. Specifically, genes for the E1 (Noc_2800), HPR (Noc_2799), an Ntr-type IIa
586 (Noc_2795), a mannose/fructose-type IIa (Noc_2313), as well as an HPr kinase/phosphorylase
587 (Noc_2796) were found. However, genes encoding components IIb, IIc, and IId were absent
588 making it unlikely that *N. oceani* expresses a functional sugar-transporting PTS system. Instead,
589 as has been seen in other Proteobacteria, these genes may form part of a regulatory cascade
590 involving RpoN (13).

591

592 ***Growth, sensing, responses and stress tolerance***

593 **Carbon fixation:** The *N. oceani* genome encodes a form I ribulose-1,5-bisphosphate
594 carboxylase/oxygenase (RuBisCO). The form I genes (*cbb*L and *cbb*S) occur in an operon with

595 *cbbX* and a hypothetical gene (Noc_0330-0333). The operon is preceded by a divergently
596 transcribed *cbbR* gene (Noc_0334) encoding a LysR-type transcriptional regulator. The *cbbX*
597 gene is required for efficient autotrophic growth in *Rhodobacter sphaeroides* (24) and is
598 predicted to be an AAA-family ATPase (which can be involved in chaperonin-like functions).
599 Both prokaryotic/plant-type (Noc_1341) and eukaryotic-type (Noc_1132) carbonic
600 anhydrases are encoded in the genome. There is no obvious candidate for a bicarbonate
601 acquisition system, however. Furthermore, the genome of *N. oceani* lacks genes for carboxysome
602 formation.
603 Genes for all enzymes to complete the Calvin-Benson-Bassham cycle are present.
604 Transketolase, NAD-dependent glyeraldehyde-3-phosphate dehydrogenase (EC:1.2.1.12),
605 phosphoglycerate kinase, pyuvate kinase, and fructose-1,6-bisphosphate aldolase are encoded by
606 an operon (Noc_2804-2808), whereas fructose-1,6-/sedoheptulose-1,7-bisphosphatase
607 (Noc_0021), ribose-5-phosphate isomerase (Noc_2667), and phosphoribulokinase (Noc_2826)
608 are encoded by isolated genes. Genes encoding ribulose-5-phosphate 3-epimerase and
609 phosphoglycolate phosphatase are grouped in an operon (Noc_2492-1493).
610 Due to the aforementioned homology of ammonia monooxygenase and particular
611 methane monooxygenase as well as the hypothesized lateral transfer of genes, the genome of *N.*
612 *oceani* was investigated for the presence of Carbon-1 (C1) metabolic pathways (see Table 4).
613 Whereas the *N. oceani* genome does not contain the genes needed to fix carbon directly from
614 methane or methanol via the RuMP or Serine pathways, it contains the genetic inventory to
615 funnel C1-carbon into the Calvin-Benson-Basham cycle (Table 4). Remarkably, two independent
616 pathways for formaldehyde oxidation to formate (Noc_1394 & Noc_1440 and Noc_2006) and
617 two gene clusters encoding the capacity for formate dehydrogenation (Noc_1122-1124;

Noc_2559-2561) were identified. The identification of this inventory provides a theoretical explanation of earlier findings that labeled C1-carbon was assimilated into the biomass of ammonia-dependently growing *N. oceani* cultures (36, 71).

Phosphorus, Nitrogen and Sulfur cycling (→ **Figure 3. Potential roles of polyphosphate and pyrophosphate in *N. oceani;* Figure 4. Schematic representation of nitrogen metabolism in *N. oceani)*:** *Nitrosococcus oceani* appears to be an efficient phosphorus sink in the oceans (Fig. 3). Some of the uptake and processing capacity for phosphorus is arranged in the *N. oceani* genome as a super cluster of genes starting with gene Noc_2388 encoding polyphosphate kinase (EC:2.7.4.1), followed by a string of 8 genes (Noc_2394-2401) encoding exopolyphosphatase (EC:3.6.1.11), regulatory protein PhoU, the high-affinity, binding protein-dependent ABC transporter PstBACS, and the two-component system PhoRB. The genome also encodes an inorganic pyrophosphatase (Noc_1134; EC:3.6.1.1, COG0221) and a V-type $H^{(+)}$-translocating pyrophosphatase (Noc_1901, COG3808). A gene homologue encoding the inorganic phosphate transporter, PitA, was not found in the *N. oceani* genome. Therefore, like in several archaea and yeasts, polyphosphate-copper complexes formed to control copper concentrations at levels above toxicity, may be removed from the cytoplasm by a phosphate:proton symport protein in the major facilitator superfamily unrelated to PitA (63). Taken together, *N. oceani* appears to have the capacity to utilize stored polyphosphate molecules as phosphagens for ATP synthesis, for substrate phosphorylation and the regulation of enzyme activity, for the direct generation of the proton motive force via pyrophosphate, and to detoxify copper (Fig. 3).

As a nitrifying bacterium, *N. oceani* affects the nitrogen cycle in its marine environment by assimilatory and dissimilatory activities (Fig. 4). While the ammonia oxidation capacity

641 (AMO, HAO; see above) of AOB link the oceanic pools of reduced (ammonia) and oxidized
642 (nitrite) nitrogen, their classical denitrification capacity through dissimilatory nitrite reductase
643 (NirK, EC:1.7.2.1, Noc_0089) and nitric oxide reductase (Nor, Noc_1847-1851) is likely a major
644 source of nitrous oxide emitted from the oceans (20, 21, 57). Additionally, there are various
645 genes that encode putative cytoplasmic (Noc_0889, Noc_2605) and exported (Noc_1542) multi-
646 copper oxidases (MCO), and other MCOs that have been implicated in oxidation of NO_x such as
647 the pan1-type MCO (Noc_0889), an alternative NO_x reductase (6). In contrast to its organization
648 in the genome of *N. europaea*, where it is clustered in a four-gene operon with the NirK-type
649 nitrite reductase (6, 16), the pan1-type MCO gene Noc_0889 is clustered, but not necessarily in
650 the same transcriptional unit, with a gene encoding cytochrome P460 (Noc_0890) directly
651 upstream of the *hao* gene cluster (7). Cytochrome P460 has been implicated in hydroxylamine
652 detoxification in several bacteria including *N. oceani* (9, 10, 31, 79). Because cytochrome P460
653 can be reduced by NO and because of the physical proximity of the two genes in the *N. oceani*
654 genome, P460 and pan1-type MCO may be components of an alternative N-oxidation pathway
655 that, in contrast to HAO, produces nitrite with NO as an intermediate. Considering the toxicity of
656 NO, and the fact that AOB have a high O_2 consumption rate but tend to live near the oxic/anoxic
657 interface, additional NO detoxification mechanisms that are operational under low O_2 conditions
658 are likely necessary for survival of AOB (75). It is thus not surprising to find in the genome of *N.*
659 *oceani* a gene, *cycP*, that encodes a beta-sheet-structured cytochrome *c*' ("c'-beta"; Noc_2696).
660 Alpha-helical cytochromes *c*' have been implicated in microaerobic NO sequestration and
661 dehydrogenation by strains of the betaproteobacterial pathogen *Neisseria* (18, 19, 52, 64, 67). It
662 has been proposed only recently that cytochrome *c*'-beta, which is evolutionarily related to and a
663 putative redox partner of cytochrome P460 (8), has evolved from an alpha-helical monoheme

664 cytochrome *c* that is ancestral to both *c*' and *c*'-beta cytochromes (M.G. Klotz and A.B. Hooper,
665 unpublished results). Both enzymes, *c*'-beta and P460, are also likely redox partners of the
666 soluble periplasmic cytochrome *c*552 (Noc_0751) as are the 3 periplasmic di-heme cytochrome *c*
667 peroxidases. The P460-heme coordination site, found in HAO and cytochrome P460, is highly
668 sensitive to hydrogen peroxide (32); therefore, it is interesting that the *c*'-beta-encoding gene is
669 clustered with one of the 3 di-heme cytochrome *c* peroxidase genes (Noc_2697), suggesting a
670 dedicated protective function of key periplasmic enzymes by this peroxidase. The genome of *N.*
671 *oceani* also contains a gene encoding the red-copper protein, nitrosocyanin (Noc_1090), a
672 putative enzyme with a cupredoxin fold (3, 55), Because this gene and the encoded protein have,
673 so far, been uniquely found in a Beta-AOB (3, 16) and now also in a Gamma-AOB, its putative
674 catalytic function is likely involved in and specific to ammonia-oxidizing catabolism. It appears
675 that the catabolic dependence on ammonia oxidation and the ultimate production of reactive and
676 toxic NO_x intermediates have imparted selective pressure on all AOB to maintain this suite of
677 periplasmic enzymes because a similar complement of genes, albeit with different genomic
678 organization, was also identified in the genome of *N. europaea* (16). It appears that the strategy
679 of maintaining a complement of NO_x-detoxifying enzymes in the periplasm is similar to the
680 strategy of active oxygen defense and designed to avoid the formation and presence of
681 nitrosating agents such as nitrous anhydride in the cytoplasm, where they have mutagenic
682 activity (75).

683 Ammonia is also the source for nitrogen assimilation and *N. oceani* has the complete
684 capacity for low (glutamate dehydrogenase) and high (glutamine synthetase; glutamine
685 oxoglutarate aminotransferase/glutamate synthase) affinity ammonia assimilation. Glutamate
686 synthase, small and large subunit-encoding genes were contiguous (Noc_1603–1604). In

addition, two genes encoding putative NADPH-ferredoxin-dependent glutamate synthase large chain proteins (Noc_2957, Noc_0101) were identified. Two NADPH specific forms of glutamate dehydrogenase were identified (Noc_2054, EC:1.4.1.3 and Noc_0864, EC:1.4.1.4). The GS-GOGAT system presumably functions at lower concentrations of ammonia. A glutamine synthetase type I (GSI, EC:6.3.1.2) is encoded by the *glnA* gene (Noc_2652). To avoid futile cycling, GSI activity is likely regulated by adenylation; the adenylyl-transferase encoded by gene *glnE* was identified (Noc_0135). Additional regulatory proteins encoded by *glnB* (Noc_0715*)* and *glnD* (Noc_0806*)* encoding a PII uridylyl-transferase (EC:2.7.7.59) were also identified. Because these genes have been identified in the *N. oceani* genome, the regulation of N-uptake is likely dependent upon the ratio of glutamine and glutamate and proceeds via adenylylation of GS and uridylylation of proteins P_I and P_{II} as it has been experimentally determined for many other Gammaproteobacteria (50).

As a member of the Chromatiaceae, *N. oceani* should participate in the sulfur cycle beyond acquiring sulfur for biosynthesis. The identified sulfate uptake capacity allows *N. oceani* to acquire and process sulfate. Sulfate reduction may proceed via sulfate adenylyltransferase (*cysND*, Noc_2288-2289; EC:2.7.7.4) to adenosine phosphosulfate (APS) via adenylylsulphate kinase (*cysC*, Noc_2482, EC:2.7.1.25) to PAPS, and via phosphoadenosine phosphosulfate (PAPS) reductase (*cysH*, Noc_2290, EC:1.8.4.8) to sulfite, which may be further reduced to H_2S by an NADPH-dependent sulfite reductase, EC:1.8.1.2 (alpha-subunit CysI: Noc_1305, beta-subunit CysI: Noc_1306). Dihydrogen sulphide is required for cysteine biosynthesis and genes encoding a thioredoxin-disulphide reductase (Noc_0345, EC:1.8.1.9) and a thiol:disulphide interchange protein (Noc_0551) were identified, as was the gene encoding thiosulfate sulfurtransferase (Noc_0593, EC:2.8.1.1). The genome of *N. oceani* also contains a gene cluster

710 that encodes a putative polysulphide reductase (*psr*, Noc_1238-1240), a monoheme cytochrome
711 (*cccA*, Noc_1241), a transporter (Noc_1242), a cytochrome *c* oxidase (Noc_1244-1247), and the
712 5 genes encoding the ubiquinone complex (Noc_1248-1252). This cluster of genes is absent from
713 the genomes of *N. europaea*, *N. eutropha*, *N. multiformis* and *N. winogradskyi* but is conserved
714 (sequence and synteny) in the genomes of *N. hamburgensis*, *Rhodopirulella baltica*, *Cytophaga*
715 *hutchinsonii* ATCC 3406 and *Cupriavidus necator* (Ralstonia eutropha JMP134). Functional
716 polysulphite reductase is a molybdopterin oxidoreductase complex that has been experimentally
717 described to act as a quinole oxidase in *Wolinella succinogenes* (41, 42). Analysis of the *N.*
718 *oceani* genome did not reveal a molybdopterin guanine dinucleotide-binding protein subunit-
719 encoding gene in the vicinity of the cluster. This putative polysulphide reductase activity awaits
720 experimental verification in *N. oceani*, which would indicate the residence of an alternative
721 catalytic center for polysulphide reduction. If present, this molybdopterin oxidoreductase could
722 theoretically also be involved in the anaerobic reduction of nitrate, chlorate, selenate or other
723 highly oxidized minerals thereby accommodating electron disposal in a microaerophilic
724 environment near the oxic/anoxic interface.

725 In contrast to many other purple sulfur bacteria, the genome of *N. oceani* lacked genes for
726 the formation of internal or external granules of sulfur compounds.

727 **Cellular growth and motility**: The genome of *N. oceani* contains almost all the typical
728 complement of genes with an identified role in cell cycle and division of other
729 Gammaproteobacteria such as *E. coli*. The genome clearly lacks the genes encoding FtsEX
730 (involved in localization and stabilization of the septal ring), FtsN and SulA. On the other hand,
731 the genome contains three genes (Noc_0272, Noc_1903, and Noc_2569) encoding proteins with
732 domains (conserved zinc-binding motif HEXXH, ATPase domain, peptidase domain) matching

733 the cell division metalloprotease, FtsH (COG0465). An alignment of all three FtsH proteins
734 showed that they differ in sequence mostly at their N- and C-termini, which flank the ATPase
735 and peptidase domains.

736 Flagellation and motility are encoded in the *N. oceani* genome in two large gene clusters
737 (Noc_2354-2378; Noc_2155-2166) and several smaller clusters (Noc_0833-0834; Noc_0124-
738 0131; NOC2052-2053 and Noc_2683-Noc_2685). The master switch operon *flhCD* was not
739 identified and is likely absent as from the genomes of other AOB. As a likely adaptation to life in
740 the open ocean, *N. oceani* appears to have only limited chemotactic capacity because just one
741 methyl-accepting chemotaxis protein (Noc_0128) of the PilJ-type was identified. In contrast, the
742 genome of *N. europaea* contained 3 MCPs (16). The presence of a sodium-driven polar flagellar
743 motor protein (MotA, Noc_0833), which can assemble with the product of the adjacent *pomB*
744 gene (Noc_0834) in addition to the usual pmf-dependent flagellar rotation mechanism, may be
745 an adaptation to *N. oceani*'s marine lifestyle.

746 **Two-component systems** **(-> Table 5. Two-component systems):** Considering the
747 reductive evolution of AOB as concluded from analysis of the *N. europaea* genome (16), the
748 genome of *N. oceani* harbored an impressive complement of complete two-component systems.
749 The genome contained 13 genes encoding histidine protein kinases (HPK) of which 12 were
750 paired (mostly succeeded) by a response regulator (RR). In addition, 1 HPK and 11 RR-encoding
751 orphaned genes were identified (Tables 4 and S3). One of the HPK genes (Noc_1756) was
752 succeeded by a tandem of 2 RR genes (Noc_1757-1758).

753 The *N. oceani* genome also contained 6 hybrid genes whose deduced protein sequence
754 contained respective HPK and RR domains (Tab. 4). One of these hybrid-HPK genes
755 (Noc_1700) was paired with a gene encoding a RR in the LuxR/FixJ family (Noc_1701) and is

adjacent to other HPK- and RR-encoding genes. This may suggest that the RR domain in the hybrid kinase has a regulatory phosphotransferase function in a phosphorylation cascade.

Stress Tolerance: In comparison with other AOB, the *N. oceani* genome contains only a limited inventory that contributes to stress tolerance in general and oxidative stress tolerance in particular. The genome encodes a heme-containing monofunctional large subunit catalase (KatE; Noc_1165) and an iron-containing superoxide dismutase (Fe-SOD; Noc_2428), both of which are supplied with iron and heme by bacterioferritin (Bfr; Noc_1411). The genome also contains genes encoding glutaredoxin (Noc_2427), thioredoxin (TRX, Noc_0603, Noc_2583) and a thioredoxin-dependent peroxide reductase (AhpC, peroxiredoxin; Noc_1307); but lacks genes for bacterioferritin-comigratory protein, NADH-peroxiredoxin reductase (AhpF), glutathione oxidoreductase and other isozymes of hydroperoxidases (KatG, KatA, Mn-Cat) and SOD (Mn-SOD, Cu/Zn-SOD). Like *N. europaea*, the genome also lacks genes for OxyR redox-autoregulatory protein, which regulates oxidative stress tolerance (KatG, AhpC, SOD), iron and zinc transport proteins (Fur, Zur) and the stationary phase-specific sigma factor, RpoS, in many bacteria. In contrast to *N. europaea*, which lacks an RpoS gene, the genome of *N. oceani* contains two genes encoding RpoS (sigma-38; Noc_0183, Noc_1702). RpoS is known to regulate hydroperoxidase (KatE) and the cell shape protein BolA (Noc_2387). *Nitrosococcus oceani* seems minimally prepared to respond to other stresses. In addition to two genes that encode the minimal growth sigma factor RpoD (sigma-70; Noc_0045, Noc_2066), the genome contains genes that encode alternative sigma factors involved in heat (RpoH, sigma-32; Noc_1935) and extreme heat (RpoE, sigma-24; Noc_2463) stresses, nitrogen starvation (RpoN, sigma-54; Noc_2793) and the need to move by flagellar motility (FliA, sigma-28; Noc_2155). Despite the absence of the flagellar master operon (*fhlCD*), FliA-dependent regulation of a

779 complete complement of flagella synthesis and chemotaxis gene clusters is aided by the presence
780 of only one methyl-accepting chemotaxis protein (MCP) of the pseudomonad PilJ-type (with
781 MA and HAMP domains; Noc_0128) whereas other MCPs in the Tar (CheM), Tsr (CheD), Tap
782 and Aer categories were absent from the genome. In addition, the *N. oceani* genome did not
783 contain genes with significant similarity to the two-component regulatory systems LasRI/RhlR
784 involved in homoserine lactone autoinducer synthesis (quorum sensing) as well as the regulation
785 of motility, virulence, starvation response and iron homeostasis in several Gammaproteobacteria.
786 The ferric uptake regulation protein (Fur; Noc_1194) regulates, for instance, ferric citrate
787 (FecIR) and ferrichrome (*fhu* operon) transport, exotoxin synthesis and the expression of
788 hydroperoxidases in many proteobacteria. Interestingly, the zinc uptake regulation protein (Zur;
789 Noc_2424) was found adjacent to a gene cluster that encodes a binding protein-dependent zinc
790 ABC transporter system (Noc_2421-2423) in the genome.
791
792 **CONCLUSIONS**
793
794 *Nitrosococcus oceani* is one of only two known ammonia-oxidizing bacteria classified as
795 Gammaproteobacteria, while the large majority of isolated ammonia-oxidizing bacteria are
796 classified as Betaproteobacteria. The genome sequence of the Betaproteobacterium *N. europaea*
797 is available and facilitates a comparison of the genes most similar between these two bacteria. Of
798 the 224 genes in *N. oceani* that were most similar to genes from *N. europaea*, 76 were classified
799 as hypothetical or proteins of unknown function. These hypothetical and unknown proteins are
800 slightly over-represented (34%) in this subset of genes relative to genes without function
801 prediction in the complete *N. oceani* (31.5%) and *N. europaea* (29%) genomes. Whereas the

802 number of genes without function and without similarity to other known genes (hypothetical
803 ORFs) is 10 times higher in the *N. europaea* (4.7%) vs. the *N. oceani* (0.57%) genome, the
804 number of genes without function but with similarity to genes in other genomes (conserved
805 hypothetical ORFs) was significantly higher in the *N. oceani* genome (30.9%) compared to *N.*
806 *europaea* (24.4%). For those genes assigned a function, the functions included central carbon
807 and nitrogen metabolism, electron transport, gene regulation, and transport. The *N. europaea*-
808 like genes were distributed throughout the *N. oceani* genome. Most of these genes were not
809 flanked by additional *N. europaea*-like genes, though in some cases two or three such genes were
810 contiguous. A cluster of *N. europaea*-like genes was present from Noc_1955 to Noc_1986 where
811 20 of these 31 genes were most similar to *N. europaea* genes. As additional genome sequences
812 become available, it will be of interest to carry out more detailed comparisons of the genes in
813 common among the ammonia-oxidizing bacteria. Of particular interest will be the
814 uncharacterized genes, some of which may encode functions unique to the use of ammonia as a
815 growth substrate whereas others maybe responsible for the difference between marine and the
816 freshwater-sediment-soil AOB in their abilities to form nitrification consortia (e.g., AOB and
817 NOB). Progress in ongoing and future research with nitrosococci will be made available at the
818 nitrosococcus project website (http://nitrosococcus.org).
819

820 **ACKNOWLEDGMENTS**

821 Sequencing was funded by the U.S. Department of Energy's Office of Biological and
822 Environmental Research and carried out primarily at the Joint Genome Institute. Finishing was
823 completed at Lawrence Livermore National Laboratory under the auspices of the U.S. DOE
824 Contract W-7405-ENG-48. Computational annotation was carried out at the Oak Ridge National

Laboratory. AFS, ATPP and MGK were supported, in part, by NSF grant EF-040621 and incentive funds provided by the University of Louisville. Our thanks go to High School student Ms. Ariella Barhen of Oak Ridge, TN, for help with the annotation. Special thanks to Drs. A.B. Hooper (UMN) and D.J. Bergmann (BHSU) and an anonymous reviewer for helpful comments.

REFERENCES

1. **Alzerreca, J. J., J. M. Norton, and M. G. Klotz.** 1999. The *amo* Operon in Marine, Ammonia-Oxidizing g-Proteobacteria. FEMS Microbiology Letters **180:**21-29.
2. **Ambrogelly, A., D. Korencic, and M. Ibba.** 2002. Functional Annotation of Class I Lysyl-tRNA Synthetase Phylogeny Indicates a Limited Role for Gene Transfer. J. Bacteriol. **184:**4594–4600.
3. **Arciero, D. M., B. S. Pierce, M. P. Hendrich, and A. B. Hooper.** 2002. Nitrosocyanin, a Red Cupredoxin-like Protein from *Nitrosomonas europaea*. Biochemistry **41:**1703-1709.
4. **Arp, D. J., and P. J. Bottomley.** 2006. Nitrifiers: More than 100 years from isolation to genome sequences. Microbe **1:**229-234.
5. **Barton, L. L.** 2005. Structural and functional relationships in prokaryotes. Springer Science +Business Media, Inc., New York.
6. **Beaumont, H. J. E., S. I. Lens, H. V. Westerhoff, and R. J. M. van Spanning.** 2005. Novel *nirK* cluster genes in *Nitrosomonas europaea* are required for NirK-dependent tolerance to nitrite. Journal of Bacteriology **187:**6849-6851.

7. **Bergmann, D. J., D. Arciero, A. B. Hooper, and M. G. Klotz.** 2005. Structure and sequence conservation of genes in the hao cluster of autotrophic ammonia-oxidizing bacteria: Evidence for their evolutionary history. Appl. Environ. Microbiol. **71:**5371-5382.

8. **Bergmann, D. J., B. Elmore, M. G. Klotz, and A. B. Hooper.** 2006. Cytochromes-P460 and -*c*', a New Class of High-Spin *c*-Cytochromes Differing from the Classic 4-Helical cytochromes *c*'. Appl. Environ. Microbiol. **submitted**.

9. **Bergmann, D. J., and A. B. Hooper.** 2003. Cytochrome P460 of *Nitrosomonas europaea.* European Journal of Biochemistry **270:**1935-1941.

10. **Bergmann, D. J., J. A. Zahn, A. B. Hooper, and A. A. DiSpirito.** 1998. Cytochrome P460 Genes from the Methanotroph *Methylococcus capsulatus* Bath. Journal of Bacteriology **180:**6440-6445.

11. **Bickle, T. A., and K. D.H.** 1993. Biology of DNA restriction. Microbiol Mol Biol Rev. **57:**434-450.

12. **Bock, E., I. Schmidt, R. Stueven, and D. Zart.** 1995. Nitrogen loss caused by denitrifying *Nitrosomonas* cells using ammonia or hydrogen as electron acceptor. Archives of Microbiology **163:**16-20.

13. **Boel, G., I. Mijakovic, A. Maze, S. Poncet, M. K. Taha, M. Larribe, E. Darbon, A. Khemiri, A. Galinier, and J. Deutscher.** 2003. Transcription regulators potentially controlled by HPr kinase/phosphorylase in Gram-negative bacteria. J. Mol. Microbiol. Biotechnol. **5:**206-215.

14. **Bothe, H., G. Jost, M. Schloter, B. B. Ward, and K. Witzel.** 2000. Molecular Analysis of Ammonia Oxidation and Denitrification in Natural Environments. FEMS Microbiology Reviews **24:**673-690.

15. **Bourniquel, A. A., and T. A. Bickle.** 2002. Complex restriction enzymes: NTP-driven molecular motors. Biochimie **84:**1047-1059.

16. **Chain, P., J. Lamerdin, F. Larimer, W. Regala, V. Lao, M. Land, L. Hauser, A. Hooper, M. Klotz, J. Norton, L. Sayavedra-Soto, D. Arciero, N. Hommes, M. Whittaker, and D. Arp.** 2003. Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. Journal of Bacteriology **185:**2759-2773.

17. **Colliver, B. B., and T. Stephenson.** 2000. Production of nitrogen oxide and dinitrogen oxide by autotrophic nitrifiers. Biotechnol. Adv. **18:**219-232.

18. **Cross, R., J. Aish, S. J. Paston, R. K. Poole, and J. W. B. Moir.** 2000. Cytochrome *c*' from Rhodobacter capsulatus Confers Increased Resistance to Nitric Oxide. J. Bacteriol. **182:**1442-1447.

19. **Cross, R., D. Lloyd, R. K. Poole, and J. W. B. Moir.** 2001. Enzymatic Removal of Nitric Oxide Catalyzed by Cytochrome *c*' in *Rhodobacter capsulatus*. J. Bacteriol. **183:**3050-3054.

20. **Dore, J. E., and D. M. Karl.** 1996. Nitrification in the euphotic zone as a source for nitrite, nitrate, and nitrous oxide at Station ALOHA. Limnology and Oceanography **41:**1619-1628.

21. **Dore, J. E., Popp, B.N., Karl, D.M., and Sansone, F.J.** 1998. A large source of atmospheric nitrous oxide from subtropical North Pacific surface waters. Nature **396:**63-66.

22. **Fani, R., E. Mori, E. Tamburini, and A. Lazcano.** 1998. Evolution of the structure and chromosomal distribution of histidine biosynthetic genes. Origins of Life and Evolution of the Biosphere **28:**555-570.

23. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenny, G. G. Sutton, W. Fitzhugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Wiedman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nquyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrman, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496-512.

24. **Gibson, J. L., and F. R. Tabita.** 1997. Analysis of the *cbbXYZ* operon in *Rhodobacter sphaeroides*. Journal of Bacteriology **179:**663-669.

25. **Goreau, T. J., W. A. Kaplan, S. C. Wofsy, M. B. McElroy, F. W. Valois, and S. W. Watson.** 1980. Production of NO2- and N2O by nitrifying bacteria at reduced concentrations of oxygen. Applied and Environmental Microbiology **40:**526-532.

26. **Gottschalk, G.** 1986. Bacterial metabolism, 2nd ed. ed. Springer-Verlag, New York.

27. **Grundmanis, V., and J. W. Murray.** 1982. Aerobic respiration in pelagic marine sediments. Geochimica et Cosmochimica Acta **46:**1101-1120.

28. **Hallam, S. J., T. J. Mincer, C. Schleper, C. M. Preston, K. Roberts, P. M. Richardson, and E. F. DeLong.** 2006. Pathways of Carbon Assimilation and Ammonia Oxidation Suggested by Environmental Genomic Analyses of Marine Crenarchaeota. PLoS Biology **4**.

29. **Holmes, A. J., A. Costello, M. E. Lidstrom, and J. C. Murrell.** 1995. Evidence that Particulate Methane Monooxygenase and Ammonia Monooxygenase may be Evolutionarily Related. FEMS Microbiology Letters **132:**203-208.

30. **Hooper, A. B.** 1969. Biochemical basis of obligate autotrophy in *Nitrosomonas europaea*. J. Bacteriol. **97:**776-779.

31. **Hooper, A. B., D. M. Arciero, D. Bergmann, and M. P. Hendrich.** 2005. The Oxidation of Ammonia as an Energy Source in Bacteria in Respiration., vol. 2. Springer, Dordrecht, the Netherlands.

32. **Hooper, A. B., and K. R. Terry.** 1977. Hydroxylamine oxidoreductase from *Nitrosomonas*: inactivation by hydrogen peroxide. Biochemistry **16:**455-459.

33. **Ibba, M., H. C. Losey, Y. Kawarabayasi, H. Kikuchi, S. Bunjun, and D. Soll.** 1999. Substrate recognition by class I lysyl-tRNA synthetases: a molecular basis for gene displacement. Proc. Natl. Acad. Sci. (USA) **96:**418-423.

34. **Jetten, M. S. M., M. Strous, K. T. van de Pas-Schoonen, J. Schalk, U. G. J. M. van Dongen, A. A. van de Graaf, S. Logemann, G. Muyzer, M. C. M. van Loosdrecht, and J. G. Kuenen.** 1998. The anaerobic oxidation of ammonium. FEMS Microbiol. Rev. **22:**421-437.

35. **Jiang, Q.-Q., and L. R. Bakken.** 1999. Nitrous Oxide Production and Methane Oxidation by Different Ammonia-Oxidizing Bacteria. Appl. Environ. Microbiol. **65:**2679-2684.

36. **Jones, R. D., and R. Y. Morita.** 1983. Methane oxidation by Nitrosococcus oceanus and Nitrosomonas europaea. Appl. Envir. Microbiol. **45:**401-410.

37. **Khmelenina, V. N., M. G. Kalyuzhnaya, V. G. Sakharovsky, N. E. Suzina, Y. A. Trotsenko, and G. Gottschalk.** 1999. Osmoadaptation in halophilic and alkaliphilic methanotrophs. Archives of Microbiology **172:**321-329.

38. **Koops, H.-P., B. Böttcher, U. C. Möller, A. Pommerening-Röser, and G. Stehr.** 1990. Description of a new species of *Nitrosococcus*. Archives of Microbiology **154:**244-248.

39. **Koper, T. E., A. F. El-Sheikh, J. M. Norton, and M. G. Klotz.** 2004. Urease-Encoding Genes in Ammonia-Oxidizing Bacteria. Appl. Environ. Microbiol. **70:**2342-2348.

40. **Kowalchuk, G. A., and J. R. Stephen.** 2001. Ammonia-oxidizing Bacteria: A Model for Molecular Microbial Ecology. Annu. Rev. Microbiol. **55:**485-529.

41. **Krafft, T., M. Bokranz, O. Klimmek, I. Schroder, F. Fahrenholz, E. Kojro, and A. Kroger.** 1992. Cloning and nucleotide sequence of the *psrA* gene of *Wolinella succinogenes* polysulphide reductase. European Journal Biochemistry **206:**503-510.

42. **Krafft, T., R. Gross, and A. Kroger.** 1995. The function of *Wolinella succinogenes psr* genes in electron transport with polysulphide as the terminal electron acceptor. European Journal Biochemistry **230:**601-606.

43. **Kwong, S. M., C. C. Yeo, D. Chuah, and C. L. Poh.** 1998. Sequence analysis of plasmid pRA2 from *Pseudomonas alcaligenes* NCIB 9867 (P25X) reveals a novel replication region. FEMS Microbiology Letters **158:**159.

44. **Kyrpides, N., R. Overbeek, and C. A. Ouzounis.** 1999. Universal Protein Families and the Functional Content of the Last Universal Common Ancestor. Journal of Molecular Evolution **49:**413.

45. **Könneke, M., A. E. Bernhard, J. R. de la Torre, C. B. Walker, J. B. Waterbury, and D. A. Stahl.** 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature **437:**543-546.

46. **Lolkema, J. S., Y. Chaban, and E. J. Boekema.** 2003. Subunit composition, structure, and distribution of bacterial V-type ATPases. J. Bioenerg. Biomembr. **35:**323-335.

47. **Lunn, J. E.** 2002. Evolution of sucrose synthesis. Plant Physiol. **128:**1490-1500.

48. **Mancinelli, R. L., and C. P. McKay.** 1988. The evolution of nitrogen cycling. Origins of Life and Evolution of the Biosphere **18:**311-325.

49. **McTavish, H., J. A. Fuchs, and A. B. Hooper.** 1993. Sequence of the Gene Coding for Ammonia Monooxygenase in Nitrosomonas europaea. Journal of Bacteriology **175:**2436-2444.

50. **Merrick, M. J., and R. A. Edwards.** 1995. Nitrogen control in bacteria. Microbiol. Rev. **59:**604-622.

51. **Min, B., J. T. Pelaschier, D. E. Graham, D. Tumbula-Hansen, and D. Söll.** 2002. Transfer RNA-dependent amino acid biosynthesis: An essential route to asparagine formation. PNAS **99:**2678–2683.

52. **Moir, J. W.** 1999. Cytochrome *c*' from *Paracoccus denitrificans*: spectroscopic studies consistent with a role for the protein in nitric oxide metabolism. Biochim. Biophys. Acta **1430:**65-72.

53. **Murray, R. G. E., and S. W. Watson.** 1962. Structure of *Nitrocystis oceanus* and comparison with *Nitrosomonas* and *Nitrobacter*. Journal of Bacteriology **89:**1594-1609.

54. **Nakamura, K., and N. Go.** 2005. Function and mol;ecular evolution of multicopper blue proteins. Cellular and Molecular Life Sciences **62:**2050-2066.

55. **Nakamura, K., T. Kawabata, K. Yura, and N. Go.** 2004. Novel types of two-domain multi-copper oxidases: possible missing links in the evolution. FEBS Letters **553:**239-244.

56. **Nevison, D. C., J. H. Butler, and J. W. Elkins.** 2003. Global distyribution of N_2O and the N_2O-AOU yield in the subsuface ocean. Global Biogeochemical Cycles **17**.

57. **Nevison, D. C., R. F. Weiss, and D. J. Erickson.** 1995. Global oceanic emissions of nitrous oxide. J. Geophysical Research **100:**15809-15820.

58. **Nicol, G. W., and C. Schleper.** 2006. Ammonia-oxidising Crenarchaeota: important players in the nitrogen cycle? Trends in Microbiology **14:**207.

59. **Norton, J. M., J. J. Alzerreca, Y. Suwa, and M. G. Klotz.** 2002. Diversity of ammonia monooxygenase operon in autotrophic ammonia-oxidizing bacteria. Arch. Microbiol. **177:**139-149.

60. **O'Donoghue, P., and Z. Luthey-Schulten.** 2003. On the evolution of structure in aminoacyl-tRNA synthetases. Microbiol. Mol. Biol. Rev. **67:**550-573.

61. **O'Mullan, G. D., and B. B. Ward.** 2005. Relationship of temporal and spatial variabilities of ammonia-oxidizing bacteria to nitrification rates in Monterey Bay, California. Applied and Environmental Microbiology **71:**697-705.

62. **Polycarpo, C., A. Ambrogelly, B. Ruan, D. Tumbula-Hansen, S. F. Ataide, R. Ishitani, S. Yokoyama, O. Nureki, M. Ibba, and D. Soll.** 2003. Activation of the

pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and class II lysyl-tRNA synthetases. Cell **12:**287-294.

63. **Remonsellez, F., A. Orell, and C. A. Jerez.** 2006. Copper tolerance of the thermoacidophilic archaeon *Sulfolobus metallicus*: possible role of polyphosphate metabolism. Microbiology %R 10.1099/mic.0.28241-0 **152:**59-66.

64. **Rock, J. D., and J. W. Moir.** 2005. Microaerobic denitrification in *Neisseria meningititis*. Biochemical Society Transactions **33:**134-136.

65. **Schmidt, I., C. Hermelink, K. van de Pas-Schoonen, M. Strous, H. J. op den Camp, J. G. Kuenen, and M. S. M. Jetten.** 2002. Anaerobic Ammonia Oxidation in the Presence of Nitrogen Oxides (NOx) by Two Different Lithotrophs. Appl. Environ. Microbiol. **68:**5351-5357.

66. **Schmidt, I., R. J. M. van Spanning, and M. S. M. Jetten.** 2004. Denitrification and ammonia oxidation by Nitrosomonas europaea wild-type, and NirK- and NorB-deficient mutants. Microbiology **150:**4107-4114.

67. **Turner, S. M., J. W. Moir, L. Griffiths, T. W. Overton, H. Smith, and J. A. Cole.** 2005. Mutational and biochemical analysis of cytochrome *c*', a nitric oxide-binding lipoprotein important for adaptation of *Neisseria gonorrhoeae* to oxygen-limited growth. Biochem. J. **388:**545-553.

68. **Umbarger, H. E.** 1978. Amino acid biosynthesis and its regulation. Ann. Rev. Biochem. **47:**533-606.

69. **Voet, D., and J. G. Voet.** 1995. Biochemistry, 2nd Edition ed. John Wiley and Sons, NewYork NY USA.

70. **Voytek, M. A., and B. B. Ward.** 1995. Detection of ammonium-oxidizing bacteria of the Beta-Subclass of the Class Proteobacteria in Aquatic samples with the PCR. Appl. Environ. Microbiol. **61:**1444-1450.

71. **Ward, B. B.** 1987. Kinetic studies on ammonia and methane oxidation by Nitrosococcus oceanus. Arch. Microbiol. **147:**126-133.

72. **Ward, B. B.** 1982. Marine ammonium-oxidizing bacteria: Abundance and activity in the Northeast Pacific Ocean. Ph.D. dissertation. University of Washington, Seattle, WA.

73. **Ward, B. B.** 2000. Nitrification and the marine nitrogen cycle, p. 427-454. *In* D. L. Kirchman (ed.), Microbial ecology of the oceans. Wiley-Liss, New York, N.Y.

74. **Ward, B. B., and G. D. O'Mullan.** 2002. Worldwide distribution of *Nitrosococcus oceani*, a marine ammonia-oxidizing gamma-proteobacterium, detected by PCR and sequencing of 16S rRNA and amoA genes. Appl Environ Microbiol **68:**4153-4157.

75. **Weiss, B.** 2006. Evidence for Mutagenesis by Nitric Oxide during Nitrate Metabolism in Escherichia coli. Journal of Bacteriology **188:**829-833.

76. **Xie, G., C. A. Bonner, and R. A. Jensen.** 1999. A probabale mixed-function supraoperon in *Pseudomonas* exhibits gene organization features of both intergenomic conservation and gene shuffling. Journal of Molecular Evolution **49:**108-121.

77. **Xie, G., N. O. Keyhani, C. A. Bonner, and R. A. Jensen.** 2003. Ancient Origin of the Tryptophan Operon and the Dynamics of Evolutionary Change. Microbiol. Mol. Biol. Rev. **67:**303-342.

78. **Zaccone, R., G. Caruso, and M. Azzaro.** 1996. Detection of *Nitrosococcus oceanus* in a Mediterranean lagoon by immunofluorescence. J. Appl. Bacteriol. **80:**611-616.

79. **Zahn, J. A., C. Duncan, and A. A. DiSpirito.** 1994. Oxidation of Hydroxylamine by Cytochrome P-460 of the Obligate Methylotroph *Methylococcus capsulatus* Bath. Journal of Bacteriology **176:**5879-5887.

80. **Zehr, J. P., and B. B. Ward.** 2002. Nitrogen Cycling in the Ocean: New Perspectives on Processes and Paradigms. Appl. Environ. Microbiol. **68:**1015-1024.

TABLES

Table 1. General features of *N. oceani* ATCC19707[#]

	Number	% of Total
DNA, total number of bases	3522111	100.00%
DNA coding number of bases	3055969	86.77%
DNA G+C number of bases	1773098	50.34%
DNA scaffolds	2	100.00%
Genes total number	3183	100.00%
Protein coding genes	3132	98.40%
RNA genes	51	1.60%
rRNA genes	6	0.19%
5S rRNA	2	0.06%
16S rRNA	2	0.06%
23S rRNA	2	0.06%
tRNA genes	45	1.41%
Other RNA genes	0	0.00%
Genes with function prediction	2130	66.92%
Genes without function prediction	1002	31.48%
Genes w/o function with similarity	984	30.91%
Genes w/o function w/o similarity	18	0.57%
Pseudo Genes	80	2.51%
Genes assigned to enzymes	364	11.44%
Genes connected to KEGG pathways	273	8.58%
Genes not connected to KEGG pathways	2859	89.82%
Genes in ortholog clusters	2901	91.14%
Genes in paralog clusters	365	11.47%
Genes in COGs	2435	76.50%
Genes in Pfam	2188	68.74%
Genes in InterPro	2065	64.88%
Genes with MyIMG Annotation	0	0.00%
Pfam clusters	1486	20.48%

[#]Derived from the DOE-JGI IMG server: http://img.jgi.doe.gov

Table 2. Mobile elements in the genome of *N. oceani* ATCC19707

[a]Repeat #	No. of ORFs in repeat	Predicted product	No. of copies	Comments and additional copies
1	2 orfs	transposases	15	
2	1 orf	transposase	3	
3	2 orfs	transposases	2	1 degenerate copy
4	1 orf	transposase	2	
5	1 orf	hypothetical (with signal peptide at N-term)	3	1 degenerate copy
6	1 orf	transposase	3	
7	1 orf	Fatty acid desaturase	2	
10	1 orf	phage integrase	2	
12	1 orf	Phosphoenolpyruvate protein phosphatase	2	only portion
13	1 orf	EF-Tu	2	
14	1 orf	Ammonia permease (most of gene length)	2	

[a]For individual listing of repeats refer to Table S1 in the online supplement.

Table 3. Transporters in the genome of *N. oceani* ATCC19707

Transporter Type[a]	Number of genes (% Total)[b]	Function
ATP-Dependent	98 (3.2%)	Substrate transport driven by ATP hydrolysis
Ion Channels	10 (0.3%)	Energy-independent facilitated diffusion
PTS System	4 (0.1%)	PEP-dependent Phosphotransferase
Secondary Transporters	78 (2.5%)	Electrochemical potential-driven transporters
Type II Secretion	15 (0.5%)	General secretory pathway
Type IV Secretion	14 (0.5%)	Conjugal DNA-Protein Transfer
Iron uptake	22 (0.7%)	Permeases and TonB-dependent Fe-siderophore receptors
Unclassified	27 (0.9%)	

[a]For individual listing of repeats refer to Table S2 in the online supplement.

[b]3093 total protein encoding genes.

Table 4. C1-carbon metabolism in the genome of *N. oceani* ATCC19707 in comparison with the genomes of the methanotroph, *M, capsulatus* (Bath) and the beta-AOB *N. europaea*.

Process	Protein	*Nitrosococcus oceani*	*Methylococcus capsulatus*
*** Methane oxidation to methanol**	AMO / pMMO (EC 1.13.12.-)	Noc_2501-2503	MCA1796-1798, MCA0295, MCA2853-2855
	Soluble methane monooxygenase (EC 1.14.13.25)	Not present	MCA1194-1205
*** Methanol oxidation to formaldehyde**	Methanol dehydrogenase cluster (EC:1.1.99.8)	Not present	MCA0299-0300, MCA0778-0789, MCA1525, 1528-1530
	Protein with PQQ repeat domain (EC:1.1.99.-)	Noc_0821	MCA2891
	Pyrrolo-quinoline quinone (PQQ) proteins	Noc_2620-2623	MCA1445-1449
*** Formaldehyde (FA) oxidation to formate**			
Glutathione (GSH)-Dependent Pathway	Dye-linked FA DH (EC:1.2.99.3)	Noc_2006	MCA2155
S-hydroxy methyl-GSH to S-formyl GSH	GSH-dependent FA DH (EC:1.2.1.1)	Noc_1394	not present
S-formyl GSH to formate	GSH S-transferase (EC:3.1.2.12)	Noc_1440	not present
Tetrahydrofolate (THF)-Dependent Pathway			
Condensation of THF with formaldehyde	5,10-methylene THF reductase (EC:1.7.99.5)	Noc_2680	MCA0137
Methylene THF to methenyl THF	Methylene THF DH (EC:1.5.99.9)	not present	MCA0508,
	Methylene THF DH (EC:1.5.1.5)	Noc_2248	MCA3018, MCA3019
Methenyl THF to formyl THF	Methenyl THF cyclohydrolase (EC:3.5.4.9)	not present	MCA0507
Formyl THF to formate	Formyl THF hydrolase (EC:3.5.1.10)	Noc_1789	not present
Formyl THF to formate	5-formyl THF cyclo-ligase (EC:6.3.3.2)	not present	MCA2773
Tetrahydromethanopterin (THMPT)-Dependent Pathway			
Condensation of THMPT with FA	Formaldehyde activating enzyme (EC:4.3.-.-)	not present	MCA2778
Methylene THMPT to methenyl THMPT	Methylene THMPT DH (EC:1.5.99.9)	not present	MCA0508,
	Methylene THMPT DH (EC:1.5.1.5)	Noc_2248	MCA3018, MCA3019
Methenyl THMPT to formyl THMPT	Methenyl THMPT cyclohydrolase (EC:3.5.4.27)	not present	MCA2863
Formyl THMPT to formate	Formyltransferase/hydrolase complex (EC:1.2.99.5 & EC:2.3.1.101)	Noc_0022-0025	MCA2857-2860
*** Formate oxidation to CO_2**	Formate dehydrogenase (EC:1.2.1.2)	Noc_1122-1124	MCA1391-1393,
		Noc_2559-2561	MCA2576-2577,
	Formate dehydrogenase (EC:1.2.1.43)	not present	MCA1208-1210
*** C_1 Assimilation**			
Ribulose Monophosphate Pathway	Hexulose 6-P synthase (EC:4.1.2.-)	not present	MCA3043, MCA3049
	Hexulose 6-P isomerase (EC:5.3.-.-)	not present	MCA3044, MCA3050
Serine Cycle	Malate dehydrogenase (EC:1.1.1.37)	not present	MCA0610
	Malyl-CoA synthetase (EC:6.2.1.9)	not present	not present
	Malyl-CoA lyase (EC: 4.1.3.24)	not present	MCA1739

Table 5[#]. Two-component systems in the genome of *N. oceani* ATCC19707

Response Regulators	25
Paired	14*
Orphaned	11
Histidine Kinases	13
Paired	12
Orphaned	1
Hybrids	6

[#]For individual listing of Two-component systems refer to Table S3 in the online supplement.

*RRs Noc_1757 and Noc_1758 both may be paired with HK Noc_1756

*HK-RR Noc_1700 may be paired with RR Noc_1701

FIGURE LEGENDS

Fig. 1. Transmission electron micrograph of *Nitrosococcus oceani* ATCC 19707. Cultures of *N. oceani* were grown until slowing readjustment of the pH with sodium carbonate indicated the beginning transition of the culture into stationary growth phase. Cells were harvested by centrifugation and sent to the University of Wisconsin-Madison Electron Microscopy facility for further processing and electron microscopy. The scale bar at the lower right (500 nm) indicates the average cell size of 1.5 µm in diameter.

Fig. 2. (A) The chromosome and (B) circular plasmid of *Nitrosococcus oceani* ATCC 19707. The outer two circles depict predicted protein-encoding and structural-RNA genes on the plus and minus strand, respectively (green, energy metabolism; red, DNA replication; magenta, transcription; yellow, translation; orange, amino acid metabolism; dark blue, carbohydrate metabolism; pale red, nucleotide metabolism; black, coenzyme metabolism; cyan, lipid metabolism; light blue, cellular processes; brown, general function; gray, hypothetical and conserved hypothetical genes; pale green, structural RNAs). Circles 3 and 4 (in panel A only) indicate on the plus and minus strand, respectively, the locations of the two *rrn* operons (black), loci involved in ammonia and urea catabolism (blue), loci involved in electron transfer and the generation of reducing equivalents (red), and predicted terminal electron acceptors (green). The two inner circles indicate GC bias and GC skew.

Fig. 3 Proposed roles of polyphosphate and pyrophosphate in the cell of *Nitrosococcus oceani*. The figure illustrates the mechanisms for phosphate uptake and the proposed flow of

phosphate, pyrophosphate and polyphosphate and their involvement in energy metabolism (ATP, pmf), central pathways (PP-Pfk) and stress tolerance (Copper detox) in the *N. oceani* cell. The proposed phosphate-proton exchanger to remove polyphosphate-copper complexes has yet to be experimentally identified.

Fig. 4 Schematic representation of nitrogen metabolism in *Nitrosococcus oceani*. The figure illustrates transport and metabolism of nitrogenous compounds and provides involved proteins identified by their gene numbers. Nitrogen is either acquired as an inorganic compound in form of ammonia/ammonium or as an organic compound in form of urea, which is hydrolyzed by urea hydrolase (UreABC) to ammonia and carbon dioxide. For synthesis, nitrogen can be assimilated from ammonia into glutamate via GDH (GdhA) or the GS-GOGAT (GlnA, GltBD) pathways. For catabolism, ammonia is oxidized by AMO and HAO and the extracted electrons are used to reduce the quinone pool via cytochromes *c*554 and *c*m552. Some of these electrons are employed to generate pmf by the cytochrome bc_1 complex and these electrons are then relayed to a terminal electron acceptor (COXaa_3, Cu-NIR, Cyt *c*', NORbb_3, etc.) via cytochrome *c*552. The question marks indicate that the active site of ammonia monooxidation by AMO needs to be still elucidated and that the annotation of AmtB-type ammonia transport genes needs experimental verification.

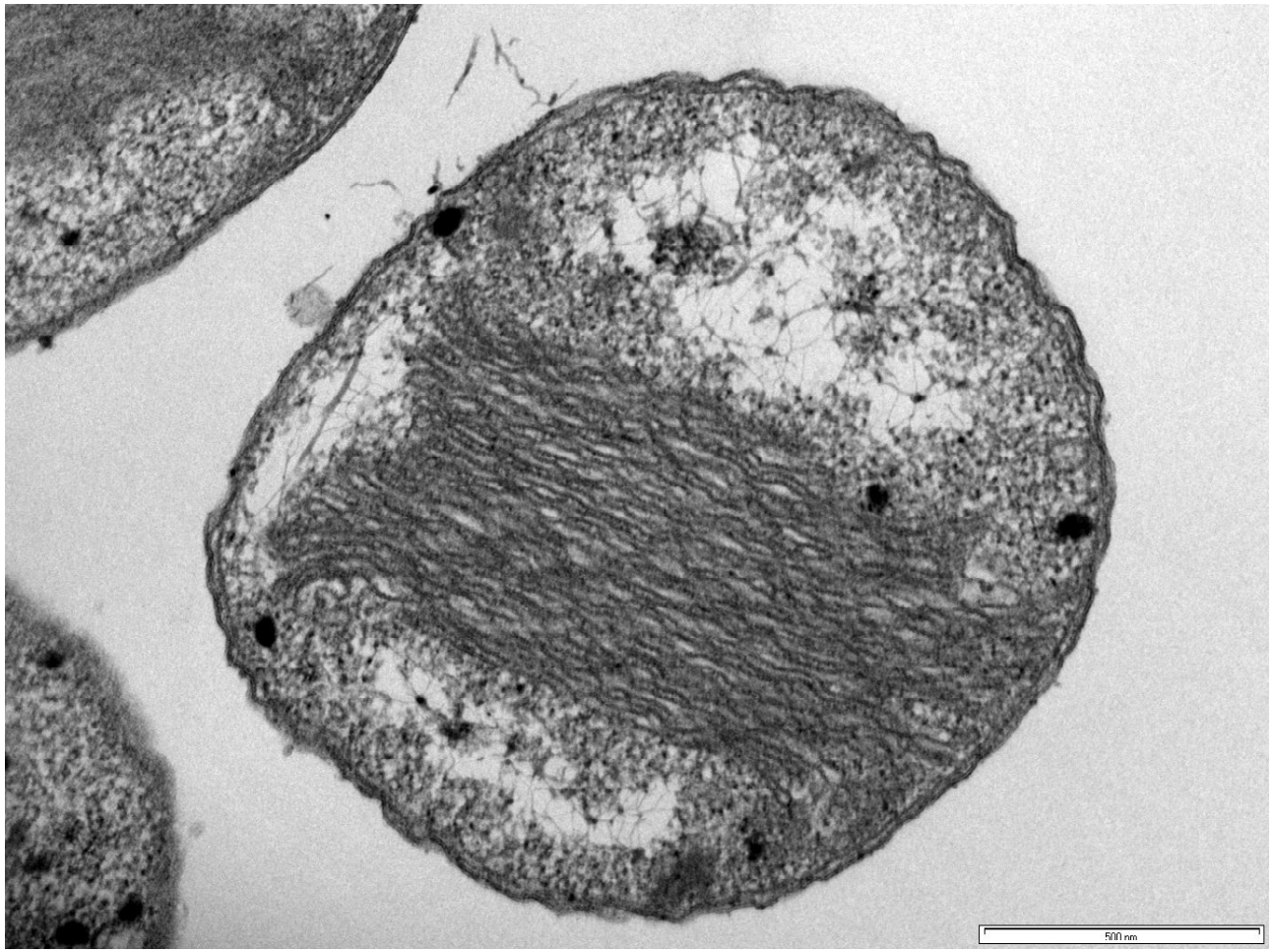

Fig. 1

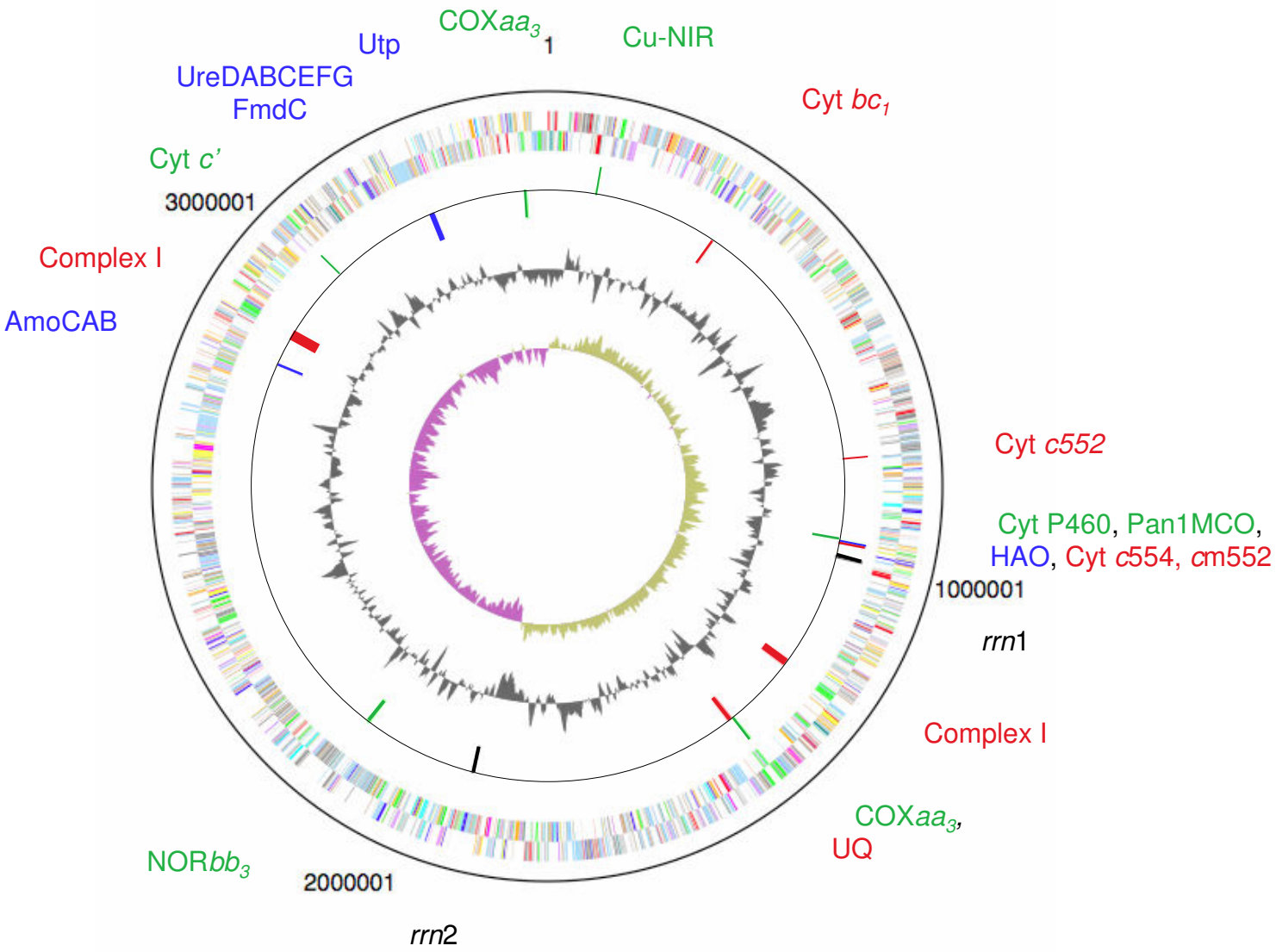
COXaa_3
1
Cu-NIR
Cyt bc_1
Utp
UreDABCEFG
FmdC
Cyt *c'*
3000001
Complex I
AmoCAB
Cyt *c552*
Cyt P460, Pan1MCO,
HAO, Cyt *c*554, *c*m552
1000001
*rrn*1
Complex I
COXaa_3,
UQ
NORbb_3
2000001
*rrn*2

Fig. 2A

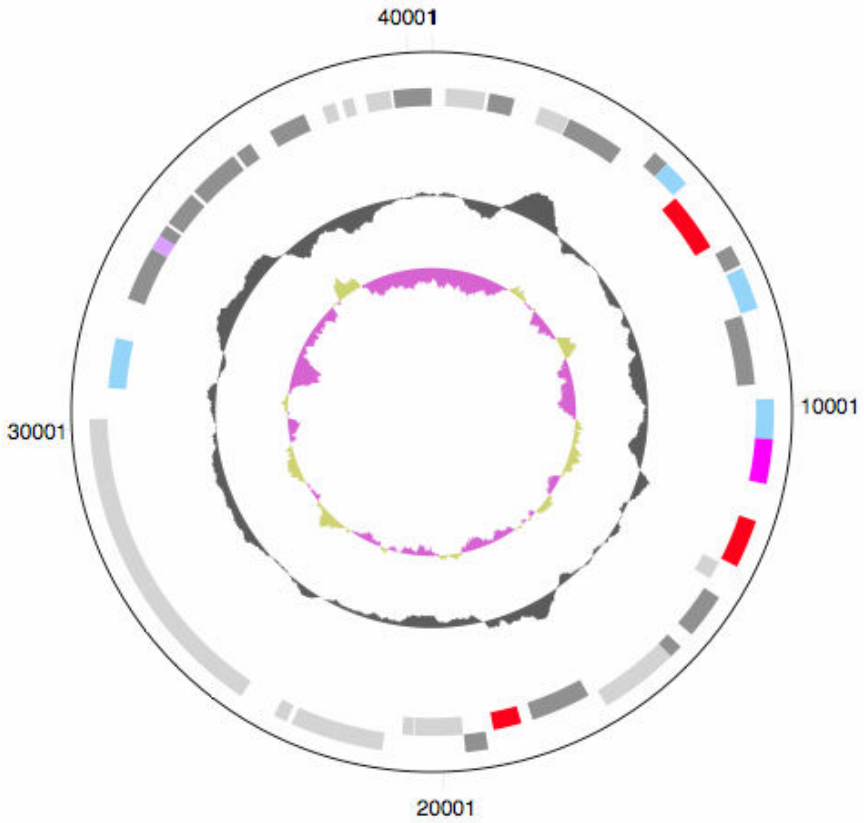

Fig. 2B

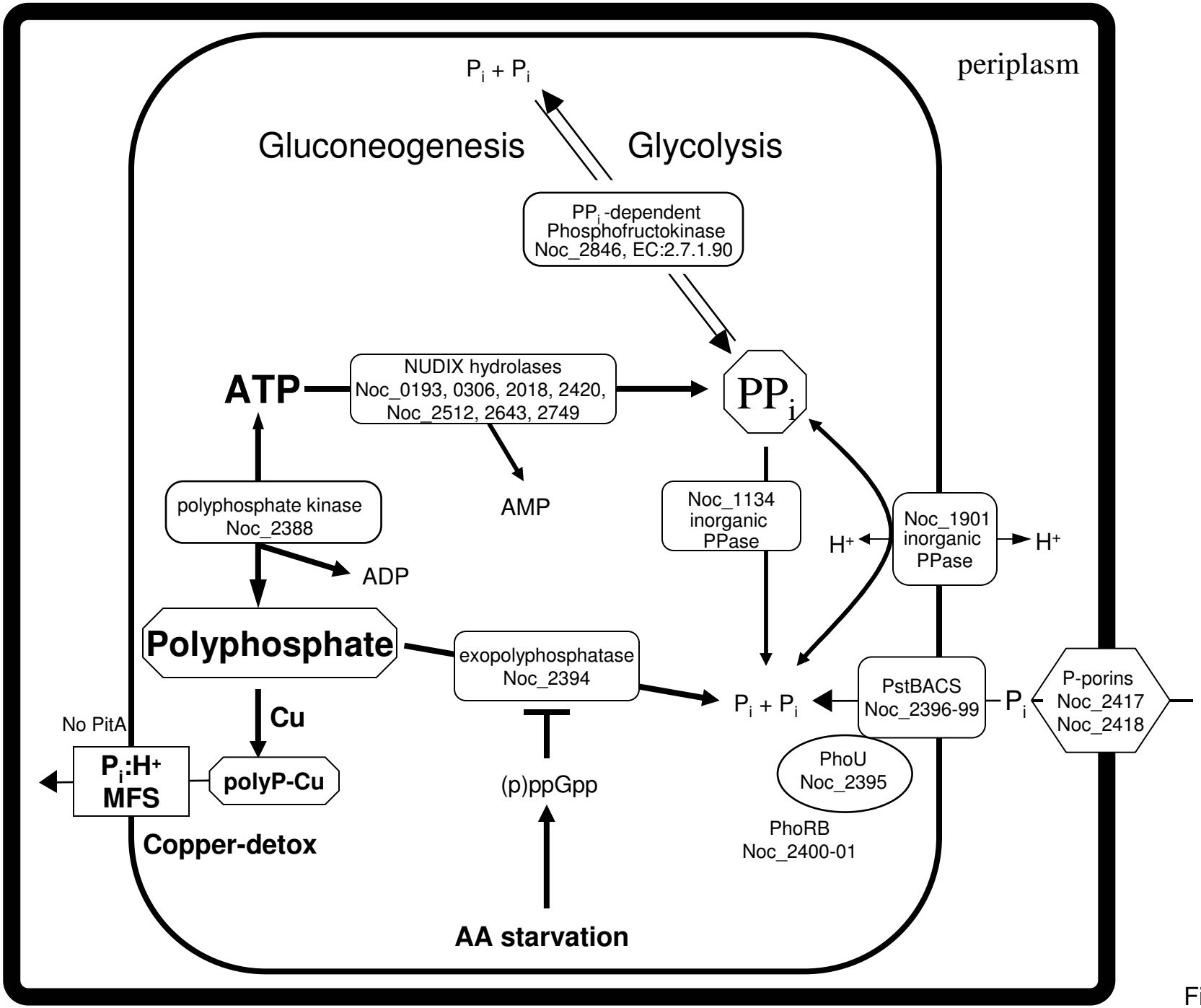
periplasm
$P_i + P_i$
Gluconeogenesis
Glycolysis
PP_i -dependent Phosphofructokinase Noc_2846, EC:2.7.1.90
ATP
NUDIX hydrolases Noc_0193, 0306, 2018, 2420, Noc_2512, 2643, 2749
PP_i
AMP
polyphosphate kinase Noc_2388
ADP
Noc_1134 inorganic PPase
H^+
Noc_1901 inorganic PPase
H^+
Polyphosphate
exopolyphosphatase Noc_2394
$P_i + P_i$
PstBACS Noc_2396-99
P_i
P-porins Noc_2417 Noc_2418
Cu
No PitA
P_i:H^+ MFS
polyP-Cu
Copper-detox
(p)ppGpp
PhoU Noc_2395
PhoRB Noc_2400-01
AA starvation

Fig. 3

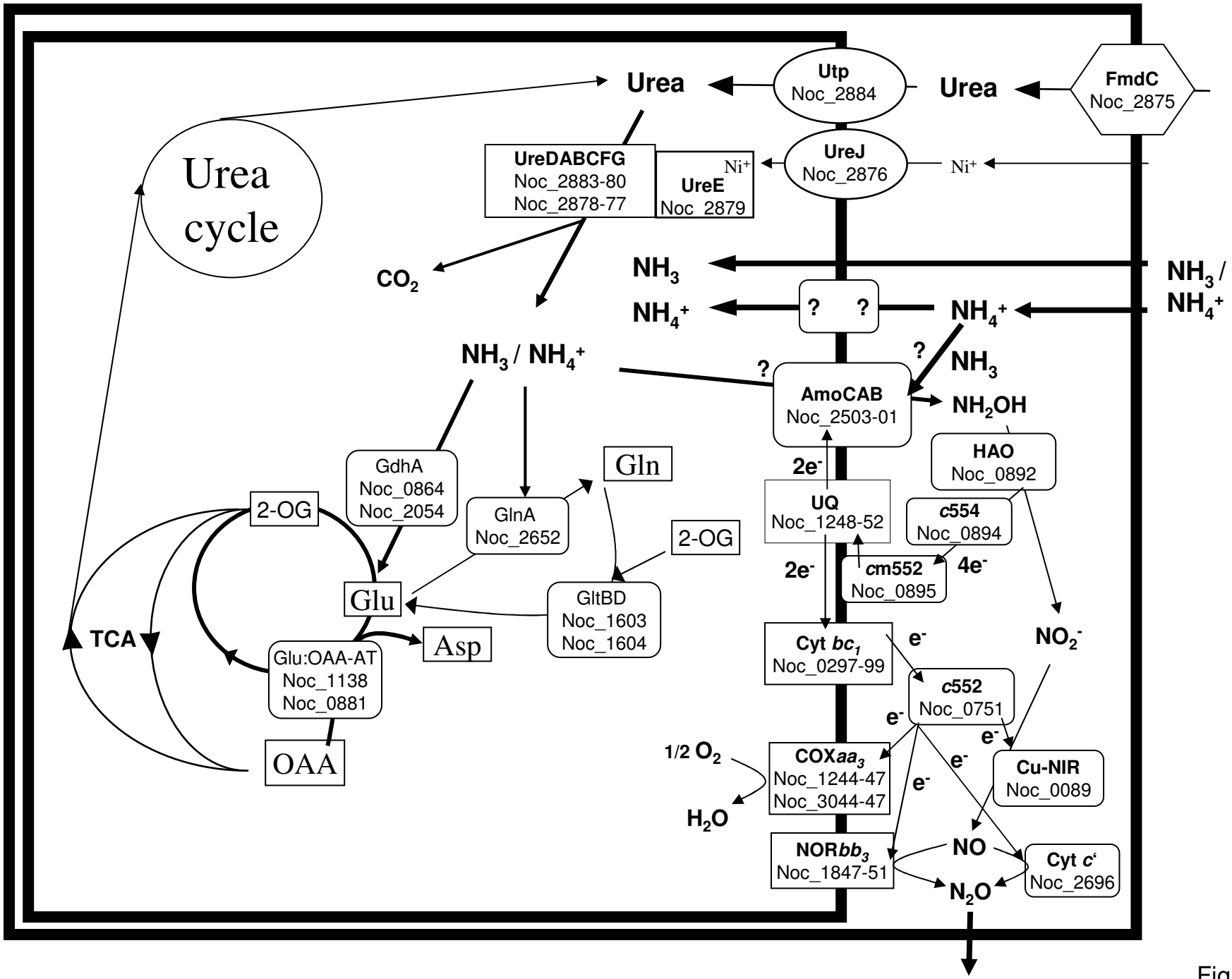
Urea
cycle
Urea
Utp
Noc_2884
Urea
FmdC
Noc_2875
UreDABCFG
Noc_2883-80
Noc_2878-77
UreE
Noc_2879
Ni+
UreJ
Noc_2876
Ni+
CO2
NH3
NH3 /
NH4+
NH4+
? ?
NH4+
NH3 / NH4+
?
?
NH3
AmoCAB
Noc_2503-01
NH2OH
HAO
Noc_0892
2e-
UQ
Noc_1248-52
c554
Noc_0894
GdhA
Noc_0864
Noc_2054
Gln
2-OG
GlnA
Noc_2652
2-OG
2e-
cm552
Noc_0895
4e-
Glu
GltBD
Noc_1603
Noc_1604
TCA
Asp
Cyt bc1
Noc_0297-99
e-
NO2-
Glu:OAA-AT
Noc_1138
Noc_0881
c552
Noc_0751
e-
e-
1/2 O2
COXaa3
Noc_1244-47
Noc_3044-47
e-
Cu-NIR
Noc_0089
H2O
NORbb3
Noc_1847-51
NO
Cyt c'
Noc_2696
N2O
OAA

Fig. 4